

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 5814 Seconds  
(without alignments)  
7273.199 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWIDFRGCCCHQN.....CEDASDTSLLGVLPASEATK 1453

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus.p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09759130/runat\_06052003\_123554\_23353/app\_query.fasta\_1.1607  
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=2000000000  
-USER=US09759130.ecgn.1.13897.erunat\_06052003\_123554\_23353 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_or.\*

21: em\_ov.\*

22: em\_pat.\*

23: em\_ph.\*

24: em\_pl.\*

25: em\_ro.\*

26: em\_sts.\*

27: em\_un.\*

28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pin.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	8109	99.6	4579	9	AF264014 Homo sapi
2	6107.5	75.0	3670	6	AX133829 Sequence
3	3779	46.4	2249	9	AK058038 Homo sapi
4	3441.5	42.3	3786	9	HSM130AC1
5	3429	42.1	3703	9	HSM130A
6	3429	42.1	4950	9	HSM130AC2
7	3396.5	41.7	3802	9	HSM130AE
8	3377	41.5	3707	6	AX281768
9	3272	40.2	4405	10	AF274883
10	3082	37.9	4475	4	BBWC11MR
11	2990	36.7	7656	9	HA243224
12	2987	36.7	7683	9	HA243212
13	2982	36.6	7667	9	HA297935
14	2980	36.6	7686	9	AF159456
c 15	2783	34.2	202252	2	AC131207
16	2447.5	30.1	7006	3	AF064259
17	2435	29.9	5802	6	A90822
18	2435	29.9	5802	6	AR193170
19	2435	29.9	5802	6	HSDMBT1
20	2246.5	27.6	6323	3	GY14953
21	2161	26.6	5108	4	AF043112
22	1963.5	24.1	6652	10	MMU37438
23	1856.5	22.8	6290	10	AB005909
24	1594.5	19.6	3769	3	AF228827
25	1586	19.5	2305	4	AF400115
c 26	1465	18.0	170882	2	AC107839
27	1442	17.7	3431	3	AF228826
28	1399	17.2	2423	4	SSSRP5
29	1341	16.5	1665	4	SSSRP3
30	1305.5	16.0	177768	2	AC131206
31	1304	16.0	3927	3	AF228824
32	1302	16.0	205331	2	AC124470
33	1277	15.7	3332	3	AF228825
34	1263	15.5	170882	2	AC107839
35	1205	14.8	2165	4	AF400122
36	1204	14.8	1316	4	SSSRP2
37	1176	14.5	161644	9	AL161645
38	1142.5	14.0	4344	6	AX402036
39	1142.5	14.0	4344	10	RNU32681
40	1142.5	14.0	4360	6	172544
41	1085.5	13.3	1519	4	SSSRP4
42	1078.5	13.3	1800	4	S78981
43	1041.5	12.8	1654	5	AB051832
44	968.5	11.9	2806	9	HA312755
c 45	916	11.3	192430	2	AC079542

ALIGNMENTS

RESULT 1



Db 326 GGACAAACCCGTGACTAGACATGGAAAAATTTGGCTTGATGATGTTTCTCTGTTATGGAAAT 385  
Qy 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140  
Db 386 GAGTCAGCTCTCTGGGAATGTCACACCCGGGAATGGGAAGCCATAACTGTTATCATGGA 445  
Qy 141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160  
Db 446 GAAGATGTTGGTGTGAACCTGTCAGGTGAAGCCAATCTGGGTTTGAGGCTAGTGGATGGA 505  
Qy 161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180  
Db 506 AACAACTCCCTGTTACGGAGAGTGGAGCTGAAATTCACAGAAAGGTGGGAACATATATGT 565  
Qy 181 AspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer 200  
Db 566 GATGATGGGTGGAACCTGGAATGCTGCTGCGGTGGTGTGCAGGCCAATAGGATGTCATCT 625  
Qy 201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220  
Db 626 TCTTTTATTTCTCTGGAGTTGTTAATAGCCCTGCTGATTTGGCCCCCATTTGGCTGGAT 685  
Qy 221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240  
Db 686 GACATTTTATGCCAGGGGAATGAGTTGGCACCTCTGGAATTCAGACATCGTGGATGGGGA 745  
Qy 241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260  
Db 746 AATCATGACTGCAGTCAACAATGAGGATGTCACATTAATTTGTTATGATAGTAGTGATCTT 805  
Qy 261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280  
Db 806 GAACAAAGGCTGTAGTGGAACTAACCGCTGATGGGAGAGTAGAGCTGAAATCCAA 865  
Qy 281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCys 300  
Db 866 GGAAGGTGGGGACCGTATGCCACATTAAGTGGAAACATGCTGCAGCTGATGCTGATGC 925  
Qy 301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320  
Db 926 AAGCAGTTGGAGTGGAAACCGCACTTCACTTCGCTGGCTGGCTGCTCAATTTGCAGTCAGG 985  
Qy 321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340  
Db 986 TCTGATGTTGATGGCTGATGGTGTCTCTGCTCCGGTAATGAATCTTTCTTTTGGAC 1045  
Qy 341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValle 360  
Db 1046 TGCAGACATTCGGGAACCGTCAATTTTGAAGTCTTCAATCAAAACGATGCTGCTGATC 1105  
Qy 361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380  
Db 1106 TGCTCAGATGGACGAGATTTGGAACTCCGACTAGCAGATGGAAGTAACAAATTTGTCAGG 1165  
Qy 381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400  
Db 1166 AGACTAGAGGTGAGAATTCATGAACACTGGTGGCAATATATGTGACCAGAACTGGAAGAAT 1225  
Qy 401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420  
Db 1226 GAACAAGCCCTCTGTGGTTGTTGTAAGCAGCTAGGATGTCGGTTCAGCGTCTTTGGCAGT 1285  
Qy 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440  
Db 1286 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGACACTGGG 1345  
Qy 441 AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg 460  
Db 1346 AATGAGTCACTCTCTGGGACTGCACATATGATGGAAAGCAAGCAACATGCTTCCGA 1405  
Qy 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480  
Db 1406 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGGCTTGTGCGG 1465

Qy 481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal 500  
Db 1466 SCTCATAGCCCTGTTATGGAGATTGGAGTGAAATACCAAGAGAGTGGGGACTGTG 1525  
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520  
Db 1526 TGTCATCAGAGATGGAGCACAAAGAAATGCACTGTTGTGTAAACAATTTGGGATGCGA 1585  
Qy 521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540  
Db 1586 AAGCCTATGATGTTGTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTGGCTG 1645  
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560  
Db 1646 GATGACCTTTCTGCAATGGAAATGACTCAATATCTGGGACTGTGAACACAGTGGATG 1705  
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580  
Db 1706 GGAAGCATAAATGTTGTACACAGAGAGATGTGATTTGAACCTGCTCAGGTGATGCAACA 1765  
Qy 581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600  
Db 1766 TGGGGCTGAGGCTGGTGGGCGGAGCAACCGCTGCTCGGAAGACTGGAGGTGTACTTT 1825  
Qy 601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal 620  
Db 1826 CAAGGACGCTGGGCGACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTGCAGCTGGTG 1885  
Qy 621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640  
Db 1886 TGTAGCCAGCTGGGACTGCCATCTTCTATCATTTGGCTGGGTCTGGGAAACGCTTCTACA 1945  
Qy 641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660  
Db 1946 GGATATGGAAAAATTTGGCTCGATGATGTTTCTCTGATGGAGATGAGTCAAGTCTCTCG 2005  
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680  
Db 2006 TCATGCAGGAACAGTGGGTGGGAAATTAATGACTGCAGTCAAGATGTTGGAGTG 2065  
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700  
Db 2066 ATCTGTCTCTGATGATCGGATATGGAGTGGAGCTGAGGCTTGTGGTGGAAAGCAGCGTGT 2125  
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720  
Db 2126 GGAAAAATGAGGTGAATGTCACAGGTCGCGGTGGGAATTTCTGTGTGCTAATGCGCTGGGA 2185  
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740  
Db 2186 ATGAACATTTGCTGAAGTGTGTTGGAGCAACTTCAATGTGGGTCTGCAATCAGGCTCTCC 2245  
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760  
Db 2246 AGAGAGCTCATTTACAGAAAGAAACATTAACATCTTAATGTGGAATTTCTGCTGCACT 2305  
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780  
Db 2306 GGAGGGGAAGCCCTCTCTGCGGATTTATAGGATGGAGTGGAGTGGAAACAGACTGCGTGTCA 2365  
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyVala 800  
Db 2366 TTAATATGAAGCAAGTTGATCTGCTCAGCCCAAGCCAGCCAGGCTGGTGGAGCT 2425  
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820  
Db 2426 GATATGCCCTGCTCTCGACGCTGTTGAAGTGAACATGCAAGACACATGCGCGCTCTGTGT 2485  
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840  
Db 2486 GATTCTGATTTCTCTCTTCTGCTGCCAATGTTGTCGACAGAAATTAACATGTTGGAGAT 2545

QY 841 AlanleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860  
DB 2546 GCCATATCTCTTCTGTGGAGATCACTTGGAAAAGGAATGCTCAACTGGCCGAA 2605  
QY 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880  
DB 2606 AAGTCCAGTGTGAAGGAGTGAACCTCACCTTGCATTTATGCCCATTTGTTCAACATCCG 2665  
QY 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspVal 900  
DB 2666 GAAGACACTTGTATCCACAGCAGAGAAAGTGGAGTTGTCTTCCCATATACAGATGTC 2725  
QY 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920  
DB 2726 CGACTTGTGAATGGCAATCCACAGTGTACGGCAAGTGGAGATCAACGTCTTGGACAC 2785  
QY 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940  
DB 2786 TGGGCTCCTACTGTGACACCCACTGGGACCCAGAAAGATGCCCGTGTCTATGACAGACAG 2845  
QY 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960  
DB 2846 CTCAGCTGGGACGCTCTCTACACACAGGAGGAAATATATATGGAGAAAGTGT 2905  
QY 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980  
DB 2906 CGTGTGGGGACACAGGTTTCAATGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2965  
QY 981 MetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSerValIleCysThr 1000  
DB 2966 ATGACAGTCTTGGAGACCTCCCTGTATPCCATGGAATACTGTCTCTGTGATCTGCACA 3025  
QY 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020  
DB 3026 GGAAGCTGACCCAGCCACTGTTCCATGCCCTCGCAATGTATCTGACCCATATTTGCT 3085  
QY 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgValAsp 1040  
DB 3086 GCAGTTCACAGAGGCGCTGTGATCTCTTAGAGCAAAACGGCTCCGCCCTAGTGGAT 3145  
QY 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060  
DB 3146 GGGACAGCGCTGTCCCGGAGAGTAGAGATCTATCACAGCGCTTCGTGGGGCACCATC 3205  
QY 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080  
DB 3206 TGTGATGACGGCTGGGACCTGAGCGATGCCAGTGTGTCTCAAAAGCTGGGCTGTGA 3265  
QY 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100  
DB 3266 GTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTACGGGCCCATCTGGCTG 3325  
QY 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120  
DB 3326 GATGACCTGACCTGCACAGGAATGGAGTCCCACTTGTGCGAGTCCCTCCCGCGCTGG 3385  
QY 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140  
DB 3386 GGGCAGCACCTGCAGGCACAGGAGGACGCGAGGGGTCTGCTGCTCAGAAATTCACAGCC 3445  
QY 1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160  
DB 3446 TTGAGGCTCTACAGTGAATGAATGAACAGAGAGCTGTGCGGAGATGGAGTCTCTAT 3505  
QY 1161 AsnGlyThrTrpGlySerValGlyArgAsnIleThrThrAlaIleAlaGlyIleVal 1180  
DB 3506 AACGGACCTGGGCGACGCTCGCAGGAGGAACATCACACAGCATAGCAGCAATTTGTG 3565  
QY 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200  
DB 3566 TGCAGGAGCTGGGCTGTGGGGAGATGGAGTTGTGAGCCCTCCCGCTTATTAAGACA 3625  
QY 1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220

DB 3626 GGCTCGTTTCATGTGGTGGATGACATTCAGTGTCTCTAAAACGCATATCTCCATATGG 3685  
QY 1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240  
DB 3686 CAGTGCCTGTCTCCCATGGGAGCGAAGATCTCCAGCACACAGACAGAGACTGGATC 3745  
QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260  
DB 3746 ACATGTGAAGATAGAATAAGAGTCCGTGGAGAGACACCCAGTGTCTCTGGGAGAGTGGAG 3805  
QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280  
DB 3806 ATCTGGCAGCAGGCTCTCTGGGACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGCG 3865  
QY 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSer 1300  
DB 3866 GAAGTGTGTGTACAGAGCTGGGCTGTGGCTCTCTGTGGCTGCCCTGAGGAGCGCTCG 3925  
QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320  
DB 3926 TTTGCCAGGGAACCTGGAACCATCTGTTGGATGACATGGGTGCAAGGAAATGAGTCA 3985  
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340  
DB 3986 TTTCTATGGGACTGTACGCCAAACCTGGGGACAGAGTGTCTGTGACACCAAGGAAGAT 4045  
QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuLeuAlaSerSerGlyHisLeu 1360  
DB 4046 GCTGGGTGAGTGTCTGTGGACAGTCTGGAATCACTGAATGCTCTCAGGTGCTTTA 4105  
QY 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380  
DB 4106 GCATTAATTTATCCAGTATCTTTGGGCTCTCTCCCGTCTGTGTTATTTCTATTCTC 4165  
QY 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArg 1400  
DB 4166 ACGTGTCCCGAGTTTCAGAAAACAAACATCTGCCCTCAGAGTTTCAACAGAGGAGG 4225  
QY 1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420  
DB 4226 GTTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACTGCCTCAAGAGAGAGGCCCA 4285  
QY 1421 HisGlyThrArgThrSerAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440  
DB 4286 CATGGACAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4345  
QY 1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453  
DB 4346 TCGCTGTGGGAGTCTTCTGCTCTGAAGCCACAAAA 4384

## RESULT 2

## AXI33829

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

AXI33829 3670 bp DNA linear PAT 15-MAY-2001  
Sequence 15 from Patent WO0119856.

AXI33829

AXI33829.1 GI:14139781

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 3670)  
Shimkets, R.A., Fernandes, E., Herrmann, J.L., Liu, X., Yang, M. and  
Boldog, F.L.

Secreted human proteins, polynucleotides encoding them and methods  
of using the same  
Patent: WO 0119856-A 15 22-MAR-2001;  
Curagen Corporation (US)  
Location/Qualifiers  
1..3670  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"



Qy	837	AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu	856
Db	1622	AACTGGGAGATGCCATATCTCTTCTGTGGGAGATCACTTTGGAAAGGGAATGTCTTA	1681
Qy	857	ThrTrpAlaGluLysPheGlnCysGluGlySerGlnThrHisLeuAlaLeuCysProfile	876
Db	1682	ACTTGGGCGGAAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTTGCATTTATGCCCATTT	1741
Qy	877	ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg	896
Db	1742	GTTCACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGAGTTGTCTGTGCCGA	1801
Qy	897	TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn	916
Db	1802	TATACAGATGTCCGACTTGTGAATGGCAAAATCCCAAGTGTACGGGCAAGTGGAGATCAAC	1861
Qy	917	ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal	936
Db	1862	GTGCTGGACACTGGGGCTCACTGTGTACACCCACTGGGACCCAGAGAATGCCGTGTT	1921
Qy	937	LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyLysTyrIleGly	956
Db	1922	CTATGCACAGACTCAGCTGTGGGACTGCTCTCTCAACACAGAGAGAAATATATTGGA	1981
Qy	957	GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu	976
Db	1982	GAAGAAGTCTTCGTGTGGGACACAGGTTTCATTCTTAGGGAATGAGTCACCTCTG	2041
Qy	977	AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer	996
Db	2042	GATAACTGTCAAAATGACAGTTCTTGAGCACCCTCCCTGATATCCATGAAATACTGTCTCT	2101
Qy	997	ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp	1016
Db	2102	GTGATCTGCACAGGAACCTTGACCCAGCCACATGTTCCATGCCTCGCAATATGATCTGAC	2161
Qy	1017	ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu	1036
Db	2162	CCATATTGTCTGAGTTCACAGGGCAGTGTGTTGATCTGCTTAGAGGACAAACGGCTC	2221
Qy	1037	ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe	1056
Db	2222	CGCTACTGATGGGACAGCCGCTGTCCGGGAGAGTAGAGATCTATCACAGCGGCTC	2281
Qy	1057	TrpGlyThrIleCysAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys	1076
Db	2282	TGGGSCACCATCTGTGATGAGGCTGGGACCTGAGCGATGCCACGTGCTGTGCAAAAG	2341
Qy	1077	LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly	1096
Db	2342	CTGGGCTGTGGAGTGGCTTCAATGCCAGGCTCTGTCTACTTTGGGAGGGGTCAAGG	2401
Qy	1097	ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro	1116
Db	2402	CCCATCTGGCTGGATGACCTGAATGCACAGGAATGGAGTCCCACTGTGGCAGTGCCT	2461
Qy	1117	SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer	1136
Db	2462	TCCCGCGCTGGGGGACAGCAGCTGCAGCACAAGGAGGACGAGGGGTCACTGTGTCA	2521
Qy	1137	GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu	1156
Db	2522	GAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAACAGGAGGAGCTGTGCTGGGAGATTG	2581
Qy	1157	GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle	1176
Db	2582	GAAGTCTTCTATAACGGGACCTGGGGCAGCGTCCGAGGAGGAACATCACCACAGCCATA	2641
Qy	1177	AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro	1196
Db	2642	GCAGGCATTTGTGCAGGCAGCTGGGCTGTGGGGAGAATGGAGTTGTCAAGCTGCCCT	2701
Qy	1197	LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis	1216

## TITLE NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 2249)

Sugano, S. and Suzuki, Y.

Direct Submission

## JOURNAL

Submitted (24-Oct-2001) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
(E-mail: cdna@leims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)

## COMMENT

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

## FEATURES

source

1..2249  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SYN00985"  
/tissue\_type="synovial membrane (knee)"  
/clone\_lib="SYN"  
/notes="cloning vector: pME18SFL3"

BASE COUNT 559 a 512 c 638 g 540 t

## ORIGIN

## Alignment Scores:

Pred. No.: 1,38e-237 Length: 2249  
Score: 3779.00 Matches: 682  
Percent Similarity: 99.71% Conservative: 2  
Best Local Similarity: 99.71% Mismatches: 2  
Query Match: 46.44% Indels: 0  
DB: 9 Gaps: 0

US-09-759-130b-381 (1-1453) x AK058038 (1-2249)

QY 770 IleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCys 789  
|||||  
Db 3 ATACGATGGAGTGGAAACAGACTGCGTGCATTAAATATGAGACAGTTGATCTGC 62  
QY 790 SerAlaHisArgGlnProArgLeuValGlyAlaAspMetProCysSerGlyArgValGlu 809  
|||||  
Db 63 TCAGCCACAGGAGCCAGCGTGGTGGAGCTGATATGCCCTGCTGACGCTGTGAA 122  
QY 810 ValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaAla 829  
|||||  
Db 123 GTGAACATGCACACATGCCGCTCTGCTGTGATTCTCTCTTCATCTGCTGCC 182  
QY 830 AsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHis 849  
|||||  
Db 183 AATGTGCTGTGCAGAGATTAATTAAGTGGAGATGCCATATCTCTTTCTGGGAGATCAC 242  
QY 850 PheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGluGlySerGluThr 869  
|||||  
Db 243 TTTGGAAAGCGGAATGCTTAACCTGGGCGGAAAGTTCAGTGTGAAGGAGTGAAC 302  
QY 870 HisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIleHisSerArgGlu 889  
|||||  
Db 303 CACCTTGTGCATTATGCCCATTTGTAACATCCGGAAGACACATTGTATCCACAGCAGAA 362  
QY 890 ValGlyValValCysSerArgTrpThrAspValArgLeuValAsnGlyLysSerGlnCys 909  
|||||  
Db 363 GTTGGAGTGTCTGTCTCCCATATACAGATGTCGACCTTGTGAATGGCAATCCAGTGT 422  
QY 910 AspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp 929  
|||||  
Db 423 GACGGCAAGTGAGATCAACGCTGTGGACACTGGGCTCACGTGTGACACCCACTGG 482  
QY 930 AspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThr 949  
|||||  
Db 483 GACCCAGAGATGCCGTGTCTATGACAGACAGCTACGCTGTGGAGTGTCTCTCAACC 542

QY 950 ThrGlyGlyLysTyrlleGlyGluArgSerValArgValTrpGlyHisArgPheHisCys 969  
|||||  
Db 543 ACAGGAGAAATATATATGGAGAAGAAGTGTCTGTGTGGGACACAGTTTCATTCG 602  
QY 970 LeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProProCys 989  
|||||  
Db 603 TTAGGGAATGAGTCACCTTCCTGGATAACTGTCAAAATGACAGTTCTTGGAGCACCTCCCTGT 662  
QY 990 IleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGlnProLeuPhePro 1009  
|||||  
Db 663 ATCCATGGAATACTGCTCTGTGATCTGCACAGAAAGCCCTGACCCAGCCACTCTTTCCA 722  
QY 1010 CysLeuAlaAsnValSerAspProTyrlleSerAlaValProGluGlySerAlaLeuIle 1029  
|||||  
Db 723 TGCCTCCAAATGATCTGACCCATATTTGTCAGTTCACAGAGGACGATGCTTTGATC 782  
QY 1030 CysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgVal 1049  
|||||  
Db 783 TGCTTAGAGACAAACGGCTCCGCTAGTGTGATGGGACAGCCGCTGTGCCGGAGAGTA 842  
QY 1050 GluIleTyrlleHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069  
|||||  
Db 843 GAGATCTATCACGACGGCTTCTGGGGACCATCTGTGATGACGCTGGACCTGAGCGAT 902  
QY 1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089  
|||||  
Db 903 GCCACGCTGTGTGTCAAAAGCTGGGCTGTGGAGTGGCTTCAATGCCACGCTCTGCT 962  
QY 1090 HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109  
|||||  
Db 963 CACTTTGGGAGGGTTCAGGGCCCATCTGGCTGATGACCTGAACCTGCACAGGAATGGAG 1022  
QY 1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129  
|||||  
Db 1023 TCCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGGACGACGACTGCAGCACAAAGAG 1082  
QY 1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrlleSerGluThrGluThr 1149  
|||||  
Db 1083 GACGAGGGGTCTATCTGCTCAGAAATTCACAGCTTGAAGCTCTACAGTGAACCTGAACA 1142  
QY 1150 GluSerCysAlaGlyArgLeuGluValPheTyrlleAsnGlyThrTrpGlySerValGlyArg 1169  
|||||  
Db 1143 GAGAGCTGTCTGGGAGATTGGAAGTCTCTATTAACGGGACCTGGGGCAGCTGGCAGG 1202  
QY 1170 ArgAsnIleThrThrAlaIleAlaGlyValCysArgGlnLeuGlyCysGlyGluAsn 1189  
|||||  
Db 1203 AGGAACATCACACAGCCATAGCAGGCAATGTGTGACAGGCACTGGGCTGTGGGAGAA 1262  
QY 1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209  
|||||  
Db 1263 GGAGTGTGACGCTCGCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGATGAC 1322  
QY 1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229  
|||||  
Db 1323 ATTCAGTGTCTTAAACGCATATCTCCATATGGCAGTGCCTGTCTGCCCATGGAGCGA 1382  
QY 1230 ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249  
|||||  
Db 1383 AGAATCTCCAGCCACAGAGAGACCTGGATCATGTGAAGATAGAAATAGAGTGGCT 1442  
QY 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleThrPheHisAlaGlySerTrpGlyThr 1269  
|||||  
Db 1443 GGAGGACACACCGAGTCTCTGGGAGAGTGGAGATCTGGCAGGAGGCTCTCTGGGGACA 1502  
QY 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289  
|||||  
Db 1503 GTGTGTGATGACTCTCTGGGACCTGGCCGAGCGGGAAGTGGTGTGTGCACGACGCTGGCTGT 1562  
QY 1290 GlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309  
|||||  
Db 1563 GGCTCTGCTGTGGCTGGCCCTGAGGGAGCTTCTGTGGCCAGGAACTGGGAACCATCTGG 1622

QY 1310 LeuAspMetArgCysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329  
DB 1623 TTGGATGACATCGGTGCAAGGAATGATCATTTCTATGGAGTGTACGCCAAACCC 1682  
QY 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349  
DB 1683 TGGGACAGAGTACTGTGGACACAGGAAGATGCTGGCGTGGAGTGTCTGGACAGTCG 1742  
QY 1350 LeuLysSerLeuAsnAlaSerGlyHisLeuAlaLeuLeuLeuSerSerIlePheGly 1369  
DB 1743 CTGAATCACTGAAGCTCTCAGCTCATTTAGGACATTTATTCAGATCATCTTGGG 1802  
QY 1370 LeuLeuLeuValLeuPheLeuPheLeuLeuTrpCysArgValGlnLysGlnLys 1389  
DB 1803 CTCCTTCTCTGCTGCTGTTATTTCTACGTGGTGGCGGAGTTCAGAAACAAAA 1862  
QY 1390 HisLeuProLeuArgValSerThrArgArgGlySerLeuGluGluAsnLeuPheHis 1409  
DB 1863 CATCTGCCCTCAGAGTTTCAACCAAGAGGAGGTTCTCTCGCGGAAATTTATTCAT 1922  
QY 1410 GluMetGluThrCysLeuLysArgGluAspProHisGlyThrArgThrSerAspAspThr 1429  
DB 1923 GAGATGGAGACCTGCTCAGAGAGAGACCCACATGGGACAGAACCTCAGATCACACC 1982  
QY 1430 ProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeuGlyValLeuProAlaSer 1449  
DB 1983 CCCAACCATGTTGTGAAGATCTAGCGACACATCGCTGTTGGGAGTCTTCTCCTCT 2042  
QY 1450 GluAlaThrLys 1453  
DB 2043 GAAGCCACAAAA 2054  
RESULT 4  
HSM130AC1 3786 bp mRNA linear PRI 22-NOV-1996  
LOCUS H.sapiens mRNA for M130 antigen cytoplasmic variant 1.  
DEFINITION Z22969  
ACCESSION Z22969.1 GI:312143  
VERSION 1  
KEYWORDS antigen; M130.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Micklem, K.K.  
Direct Submission  
Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department  
of Pathology, University of Oxford, Level 1, Maternity Block, John  
Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom  
2 (bases 1 to 3786)  
Law, S.K., Micklem, K.J., Shaw, J.M., Zhang, X.P., Dong, Y., Willis, A.C.  
and Mason, D.Y.  
A new macrophage differentiation antigen which is a member of the  
scavenger receptor superfamily  
Eur. J. Immunol. 23 (9), 2320-2325 (1993)  
93380506  
8370408  
Location/Qualifiers  
1. 3786  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="Macrophage"  
102..3557  
/codon\_start=1  
/product="M130 antigen cytoplasmic variant 1"  
/protein\_id="CAA80542.1"  
/db\_xref="GI:312144"  
/db\_xref="SPTREMBL:Q07899"  
/translation="MVLLDSGSADFRRHFNVLSPFTITVVLLLSACFVTSLSGTDK  
ELRLVDGNKSGRVEYVQEWGTCVNNNSMEAVSYICNLQCPPTAKAPGWNANS  
AGSGRIWHDYSCRNESALMDKHDGKHSNCTHOODAGVTCSDGNSLRLRTRGG  
NMCSGRIELKFGRWGTVCDNFNDHASVICROLECGSAVSFSGNSFNFGSGPIWF  
DDLICNGNESALWNCKHOGWKHNDHAEDAGVICSKGADLSRLVDVGTCSGRLEV

RFOEWGTCDDGDWSDAAYVACKQLCQPTAVTAIGRVNASKGFGHILWLSVSCQGEH  
PAYWCKHHEWGRHYCNHEDAGVTCSDGSLRLRLRGGRSGRAGTVEIQLLGV  
CDRGKLEADVVCRLQCGCSALKTSQVYSKIQTATWLFSSCNGNETSLMDCKNW  
QWGLTCDHYEAKITCSAHREPRVGGDIPCSGRVEVKGHDWGSICDSDFSLAAS  
VLRELQCTVVSITLGAHFGEGNQLWAEFOCEGESHLSLCPVAPRPGTCSHSR  
DGVVCSRYTEIRLVNKGTPCEGRVELKTLGAWGSLCNHWDIEDAHLVQCLQACGVA  
LSTPGGARFGKNGQIWRHMFCHTGTQEHMGDCPVTAALGASLCFSEQVASYICSGNS  
QTLSSCNSSSLGPTPTPIPEESAVACIESQGLRLVNGGRCAGRVEIYHSGSWGTICD  
DSWLDSDAHVVCRLQCGEATNATGSAHFGEGTGTPIWLDKMKCNKESRIWQCHSHG  
GQONCRHKEDAGVTCSEFMSLRITSEAREACAGRLEVFNANGTGVKSKMSSTTVG  
VVCRLGCDKGINPASLDKAMSI PMWVNDVOCPKGPDLDLDDAQVVCQCGCPALKE  
ETWITCDNKIRLOBGPTSCSGRVEIWHGSGKTVCCDDSWLDLDDAQVVCQCGCPALKE  
AFRAEEFGGTGPIWLNVEKCKGNESLWDCPPARWGHSCGKHEDAAVNCITDLSVQK  
TPKATTSRSGRSSSEFVAVGILGVLLAIFVFLFKRQRORLAVSSRGENLVHQ  
IQYREMNNSCLNADDLMLNSENSESHESADFAELISVSKFLPISGMEKEAILSTHEK  
ENGNI"  
102..221  
222..3554  
/product="M130 antigen cytoplasmic variant 1"  
/evidence="experimental"  
polyA\_signal 3760..3765  
polyA\_signal 3764..3769  
polyA\_signal 3768..3773  
polyA\_site 3783  
BASE COUNT 1006 a 746 c 1074 g 960 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,77e-215 Length: 3786  
Score: 3441.50 Matches: 608  
Percent Similarity: 70.24% Conservatives: 166  
Best Local Similarity: 55.17% Mismatches: 311  
Query Match: 42.29% Indels: 17  
DB: 9 Gaps: 6  
US-09-759-130B-381 (1-1453) x HSM130AC1 (1-3786)  
QY 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383  
DB 222 GGAACAGACAGAGAGCTGAGCTAGTGTGTAACAAAGTGTAGCGGGAGAGTGGAA 281  
QY 384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpIleAsnGlnAla 403  
DB 282 GTGAAGTCCAGAGAGAGTGGGGAACGGTGTCTAATAATGGCTGGAGCATGCAAGCGTC 341  
QY 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423  
DB 342 TCTGTGATTTGTAACAGCTGGGATGTCCAACTGTATCAAGAGCCCTGGATGGCTAAT 401  
QY 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443  
DB 402 TCCAGTGCAGGTTCTGGAGCGCATTTGGATGATCATGTTCTTCGCTGGGAGTGAATCA 461  
QY 444 AlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThr---CysPheArgArgSer 462  
DB 462 GCTCTTTGGATTGCAACATCATGATGGATGGGGAACATAGTACTGTACTCACCACAA 521  
QY 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482  
DB 522 GATGCTGGAGTACCTGCTCAGATGATGATCAATTTGGAATGAGGCTGAGCGGTGAGGG 581  
QY 483 SerProCysThrGlyArgLeuGluValLysThrGlnGlyLeuTrpGlyThrValCysHis 502  
DB 582 AATATGTGTTCTGGAGAAATAGATCAATTTCCAAAGGACGGTGGGGAACAGTGTGTGAT 641  
QY 503 AspArgTrpSerThrArgAsnAlaAlaValCysLysGlnLeuGlyCysLysPro 522  
DB 642 GATACTTCAACATAGATCATGCTGCTCATTTGTAGACAACCTGAATGTGGAAGTGTCT 701  
QY 523 MethHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleThrLeuAspAsp 542  
DB 702 GTCAGTTTCTCTGTCATCTAATTTTGGGAAGGCTGTGGACCAATCTCTGGTTGTGAT 761

QY 543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562  
Db 762 CTTATATGCAACGGAATAGTACAGCTCTGGAACCTGCAACATCAAGGATGGGAAAG 821  
QY 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582  
Db 822 CATAACTGTGATCATGCTGAGGATGCTGGAGTGTATTTGCTCAAGGAGGAGATCTGAGC 881  
QY 583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602  
Db 882 CTGAGACTGGTAGAGGACTGAAATGTTCAAGAAAGATTAGAGTGAATCCCAAGGA 941  
QY 603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSer 622  
Db 942 GAATGGGGACAATATGATGACGGCTGGACAGATTACGATGCTGCTGGCATGCAAG 1001  
QY 623 GlnLeuAspCysProSerSerIleIleGlyMetClyLeuGlyAsnAlaSerThrGlyTyr 642  
Db 1002 CAACCTGGGATGTCACACTGGCGTCACAGCCATTGGTCGAGTTAACGCCAGTAAGGGATT 1061  
QY 643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662  
Db 1062 GGACATCTGGCTTGACAGCGTTCTTGGCAGGACATGAACCTGCTGCTGGCAATGT 1121  
QY 663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682  
Db 1122 AAACACCATGAATGGGAAAGCAATTATTGCAATCACAATGAAGATGCTGGCGTGACATGT 1181  
QY 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerArgCysAlaGlyLys 702  
Db 1182 TCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGCGCCGCTGGCTGGGACA 1241  
QY 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuGlyCysAlaAsnGlyTrpGlyMetAsn 722  
Db 1242 GTTCAGCTGGAGATTCAGAGACTCTTAGGGAAGGTGTGTACAGAGCTGGGACTGAA 1301  
QY 723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742  
Db 1302 GAAGCTGATGTGGTTGTCAGGACGCTGGGATGGATCTGCACATCAAAACATCTATCAA 1361  
QY 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762  
Db 1362 GTGTACTCCAAATCCAGGCAACAAACACATGCGTGTCTTAAGTAGCTGTAAACGGAAAT 1421  
QY 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782  
Db 1422 GAACTTCTCTTGGGACTGCAAGAACTGGCAATGGGGTGACACTACTGTGATCACTAT 1481  
QY 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802  
Db 1482 GAAGAAGCCAAATTAACCTGCTCAGCCACACAGGAAACCCAGACTGGTTGGAGGGGACAT 1541  
QY 803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822  
Db 1542 CCCGTGTTGGAGCTGTGAAGTGAAGCACTGGTGACAGCTGGGCTCCATCTGTGATTCG 1601  
QY 823 AspPheSerLeuHisAlaAlaAsnValLeuGlyCysArgGluLeuAsnGlyAspAlaIle 842  
Db 1602 GACTTCTCTGGAAGCTGCCAGCTTCTATGACAGGAAATACAGTGTGGCAGACATGTC 1661  
QY 843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862  
Db 1662 TCTATCTGGGGGAGCTCACTTTGGAGAGGAAATGACAGATCTGGGCTGAAGAAATTC 1721  
QY 863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882  
Db 1722 CAGTGTGAGGACATGATGCTCCATCTTCACTCTGCCAGTAGCAGCCGCCGCCAGAGGA 1781  
QY 883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu 902  
Db 1782 ACTTGTAGCCACACAGGGATGTTGGAGTAGTCTGCTCAGATACACAGAAATTCGTTG 1841  
QY 903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922

Db 1842 GTGAATGGCAAGACCCCGTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGA 1901  
QY 923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGluLeuSer 942  
Db 1902 TCCCTCTGTAACTCTCACTGGGACATAGAAGATCCCATGTTCTTGGCAGACGCTTAA 1961  
QY 943 CysGlyThrAlaLeuSerThrGlyGlyLysTyrIleGlyGluArgSerValArgVal 962  
Db 1962 TGTGGAGTTGCCCTTCTACCCAGGAGGACGACGTTTGGAAAAGGAAATGGTCAGATC 2021  
QY 963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnAsnGlnMetThr 982  
Db 2022 TGGAGGCATATGTTTCACTGGCAGTGGGACATGAGCAGACATGGGAGATGTCTCTTA 2081  
QY 983 ValLeuGlyAlaProProCysIleHisCysAsnThrValSerValIleCysThrGlySer 1002  
Db 2082 GCTCTAGTGTCTCATTTGCTCTCAGAGCAAGTGGCCTCTGTAATCTGCTCAGGAAC 2141  
QY 1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal 1022  
Db 2142 CAGTCCCAACACACTGCTCTGTCGCAATTCATCGTCTTTGGGCCCAACAGCCCTACCAT 2201  
QY 1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042  
Db 2202 CCAGAAAGAAAGTGTGTGGCCTGCATAGAGAGTGTCAACTTCGCTGGTAAATGGAGGA 2261  
QY 1043 SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp 1062  
Db 2262 GGTCTGCTGTGGGAGTAGAGATCTATCATGAGGCTCTCTGGGGACCATCTGTGTAT 2321  
QY 1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082  
Db 2322 GACAGCTGGGACCTGAGTGTGTCACCGCTGTTGGCAGACAGCTGGGCTGTGGAGAGCC 2381  
QY 1083 PheAsnAlaThrValSerAlaHisPheClyGluGlySerGlyProIleTrpLeuAspAsp 1102  
Db 2382 ATTAATGCCACTGTTCTGCTCATTTTGGGAAGAAAGAGGAGGAGGAGGAGGAGGAG 2441  
QY 1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122  
Db 2442 ATGAATGCAATGGAAGAAAGAAATCCGCAATTTGGCAGTCCCATTTACAGCGCTGGGGGAG 2501  
QY 1123 HisAspCysArgHisLysLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142  
Db 2502 CAAATTCAGGACACAGAGGAGTGGGAGTATCTCTCAGAAATTCATGCTCTGAGA 2561  
QY 1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162  
Db 2562 CTGACCACTGAAGCCAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2621  
QY 1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182  
Db 2622 GCTTGGGGCACTGTGGCAAGAGTAGCATGTCTGAACCACTGGGTGGTGGTGGTGGAGG 2681  
QY 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202  
Db 2682 CAGCTGGGCTGTGCAGACAAAGGAAATCAACCTGCATCTTTAGACAGGCCATGTCC 2741  
QY 1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222  
Db 2742 ATTCCTCATGTGGGAGCAATGTTCAAGTGTCCAAAGAGGAGGAGGAGGAGGAGGAGGAG 2801  
QY 1223 LeuSerAlaProTrpGluArgArgIleSerProAlaGluGluThrTrpIleThrCys 1242  
Db 2802 CCATCATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGGAGGAGGAGGAGGAGGAG 2861  
QY 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262  
Db 2862 GACAACAGATAGAGACTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2921  
QY 1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluVal 1282



Qy	863	GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp	887
Db	1722	CAGTGTGAGGACATGAGTCCCATCTTTTTCACCTGTCCAGATAGCACCCGCCGAGAAGGA	1781
Qy	883	ThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeu	902
Db	1782	ACTTGTAGCCACAGAGGAGTGTGGAGTAGTCTGCTCAAGATACACAAATTCGCTTG	1841
Qy	903	ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly	922
Db	1842	GTGAATGGCAAGACCCCGTGTAGGCGCAGAGTGGAGTCTCAAAACGCTTGGTCCCTGGGGA	1901
Qy	923	SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer	942
Db	1902	TCCCTCTGTAACTCTCACTGGGACATAGAAAGATGCCCATGTCTTTTGGCAGCAGCTTAA	1961
Qy	943	CysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgVal	962
Db	1962	TGTGGAGTTCGCTTCTACCCAGAGGAGCAGCGTTTGGAAAAGAAATGGTCGAGTC	2021
Qy	963	TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr	982
Db	2022	TGGAGGCATATGTTTCACTGCCTGGACTGAGCAGCACATGGAGATGTCTCTTA	2081
Qy	983	ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer	1002
Db	2082	GCTCTAGTGCTTCAATTATGCTTCCAGAGCAAGTGGCCTCTGTAACTCTGCTCAGAAAC	2141
Qy	1003	LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaVal	1022
Db	2142	CAGTCCCAACACTGTCCTCGTGCATTCATCGCTTTGGGCCCAACAGGCCCTACCAAT	2201
Qy	1023	ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp	1042
Db	2202	CCAGAACAAAGTGTGTGGCCTCCATAGAGAGTGTCAACTTCGCTGTTAATGGAGCA	2261
Qy	1043	SerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAsp	1062
Db	2262	GGTCGCTGTCTGGAGAGTAGAGATCTATCATGAGGGCTCTCTGGGCCACCATCTGTGAT	2321
Qy	1063	AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla	1082
Db	2322	GACAGCTGGACCTGAGTGATGCCACAGTGGTTTGCAGACACTGGGTGTGGAGAGGCC	2381
Qy	1083	PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp	1102
Db	2382	ATTAATGCCACTGGTCTGCTCATTTTGGGAAGAAACAGGCCCATCTGGCTGGATGAG	2441
Qy	1103	IleuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln	1122
Db	2442	ATGAATGCAATGAAAAGAAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGCAG	2501
Qy	1123	HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg	1142
Db	2502	CAAAATTCAGGCACAGGAGGATCGGAGGATTATCTGCTCAGAATTCATGCTCTGAGA	2561
Qy	1143	IleuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuLysValPheTyrAsnGly	1162
Db	2562	CTGACCACTGAAGCCAGACAGAGCGCTGTGACGGCGTCTGGAAGTTTTTACAAATGGA	2621
Qy	1163	ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg	1182
Db	2622	GCTTGGGSCACTGTGGCAAGAGTAGCATCTCTGAAACCACTGTGGTGTGGTGGCAGG	2681
Qy	1183	GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer	1202
Db	2682	CAGCTGGGCTGTGCACACAAGGGAATAACACCCCTGCATCTTTAGACAAGGCCATGTCC	2741
Qy	1203	GlyPheMetTrpValAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys	1222
Db	2742	ATTCCTCCATGTGGTGACAATGTTCAGTGTCAAAAGGACCTGCACACGCTGTGGCAGTGC	2801
Qy	1223	IleuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys	1242

```
Db 2802 CCATCATCTCCATGGAGAGAGACTGCCAGCCCTCGGAGGACCTGGATCATCATGT 2861
Qy 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
Db 2862 GACACACAGATAAGACTTCAGAGAGACCCACTTCTGTTGGAGCTGTGGAGATCTGG 2921
Qy 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluVal 1282
Db 2922 CATGGAGTCTCTGGGGACAGTGTGATGACTCTTGGAGCTGTGGAGATCTGGAGTG 2981
Qy 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGly 1302
Db 2982 GTGTGTCAACAATCTGGTGTGGTCCAGCTTTGAAGCATTCACAAAGACGAGATTGTG 3041
Qy 1303 GlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeu 1322
Db 3042 CAGGGAGCTGGACCATATGGCTCAATGAAGTGAAGTCAAGAGGAATCAGTCTTCITG 3101
Qy 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
Db 3102 TGGGATTCCTCTCCAGACAGCTGGGGCCATAGTAGTGTGGCACAAGAGACGCTGCA 3161
Qy 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356
Db 3162 GTGAATTCACAGATATTTCAGTCGAGAAACCCACAAAAGCCACACAGGTCGCTCA 3221
Qy 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376
Db 3222 TCCCGTCAGTCATCTTTATTTCAGTCAGTGGATCCCTGGGTGTCTGTGGCCATTTC 3281
Qy 1377 Ile-----LeuPheIleuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
Db 3282 GTCCGATTATTCTTCTGACTAAAAGCGACAGACAGACGCG-----CTTGCA 3332
Qy 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
Db 3333 GTTCTCTCAAGGAGAGAACTTAGTCCACCAATTAATCCGGGAGATGAATCTTTCG 3392
Qy 1415 LeuLysArgGluAsp 1419
Db 3393 CTGAATGCAGATGAT 3407
RESULT 6
HSM130AC2 HSM130AC2 4950 bp mRNA: linear PRI 22-NOV-1996
LOCUS H.sapiens mRNA for M130 antigen cytoplasmic variant 2.
DEFINITION H.sapiens mRNA for M130 antigen cytoplasmic variant 2.
ACCESSION Z22970
VERSION Z22970.1 GI:312145
KEYWORDS antigen; antigen M130.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Micklem,K.K.
Direct Submission
Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
of Pathology, University of Oxford, Level 1, Maternity Block, John
Radcliffe Hospital, Headington, Oxford, OX3 9DU, United Kingdom
2 (bases 1 to 4950)
Law,S.K., Micklem,K.J., Shaw,J.M., Zhang,X.P., Dong,Y., Willis,A.C.
and Mason,D.Y.
A new macrophage differentiation antigen which is a member of the
scavenger receptor superfamily
Eur. J. Immunol. 23 (9), 2320-2325 (1993)
JOURNAL 93380506
MEDLINE /db_xref="taxon:9606"
PUBMED /cell_type="Macrophage"
8370408
FEATURES
Location/Qualifiers
1..4950
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="Macrophage"
```

```
CDS
102..3572
/codon_start=1
/product="M130 antigen cytoplasmic variant 2"
/protein_id="CAA80543.1"
/db_xref="GI:312146"
/db_xref="SPTREMBL:Q07900"
/translation="MVLEDSGSADFRHRHFNLSPTTITVVLALLSACFVTSLSGGTDK
ELRLVDGENKCSGRVEYKQVEWGTVCNNGWSMEAVSYICNOLGCGPTAIKAPGWANSS
AGSGRIWHDHVSRCRGNESALWDKHDGKHSNCTHOODAGVTCSDGNSLENRLRFRGG
NMGSGRIEIKFQGRWGTVCDDNFNDHSAVICROLECGSAVSFCSSNFGESGPIWF
DDLICNGESALWNCKHOGWKHCNDHAEDAGVICSKGADLSRLHVDGVTGCGSGLV
RFQEWGTCIDDDSDYDAACVACQLGCTPTAVTAIGRYNASKGFGHILWDSVSCGHE
PFWCKRHHWGHKHCNHDAGVTCSDGSDLELRGGRCAGTVEYETQRLGKV
CDGWLKEDADVVCRLGCSALTSYQVYSKIQTATWLFSSCNGNETSLWDCKNW
QMGGLTCDHYEEAKITCSAHPRLVGGDIPCSGRVEYHGDVTGSI CDSDFSLAAS
VVCRLGCTVVSILGGAHFGEGNQIWAEEFCQEGHESHSLSLCPVAPRPGTCSHSR
DVGVCYSRYTIRLVNGTKPCGRVELKTLGANGSLCHSHWDIEDAHLVLCQKCGVA
LSTGGARFGKNGOILWPHMFCTGTQEHMGDCPVTALGASLCPSEOVASVICSNGS
QTLSSCNSSSLGPTPTIPESAVACIESGQLRLVNGGRCAGRVEIHEGWSGTICD
DWSLSDAHVVCRLGCGEAINATGSAHFEGTGPIWDEMKCNKESRIWOCHSHGTV
GQOCHRKEDAGVICSSEPMISRLTSEAREACAGRLEFYINGAMGTGKSSMSETTVG
VVCRLGCDKADGKINPALSADKAMSTPMVDNVQCPKPDTLWQCPSSEWERLAPSE
ETWITCDNKIRLOEPTSCSGRVEIWHGSGWTVCDSDWLDLDDAOVVCQQLCGPALK
AFKEATGQGTGPIWLNKVKCKGNESSLWDCPAPWGHGECGHEKEDAVNCTDLSVOK
TPQKATGRRSRQSFIAVGLVGLVLLAIFVALFELTKRQRQRQLAYSSRGLENLVHQ
IQRREMSNCLNADDDLLMNSSLGVLGSIAGGFRSVAEVAQTFFDFDKLKKKNVI
GSLDAYNGQE"
102..221
222..3569
/product="M130 antigen cytoplasmic variant 2"
/revidence=experimental
polyA_signal 4924..4929
polyA_signal 4928..4933
polyA_signal 4932..4937
polyA_site 4947
BASE COUNT 1348 a 970 c 1277 g 1355 t
ORIGIN
Alignment Scores:
Pred. No.: 2,5e-214 Length: 4950
Score: 3429.00 Matches: 598
Percent Similarity: 71.27% Conservative: 161
Best Local Similarity: 56.15% Mismatches: 294
Query Match: 42.14% Indels: 12
DB: 9 Gaps: 4
US-09-759-130B-381 (1-1453) x HSM130AC2 (1-4950)
Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db 222 GGAACAGACAGAGCTAGGCTAGGTGTGTAAGCAAGTGTAGCGGAGAGTGGAA 281
Qy 384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db 282 GTGAAGTCCAGGAGAGTGGGGAACGGTGTGTATAATGCGTGGAGCATGGAAGCGGTC 341
Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db 342 TCTGTGATTGTGAACACAGCTGGATGTCCAACTGCTATCAAAAGCCCTGGATGGGCTAAT 401
Qy 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db 402 TCCAGTCGAGGTTCGTGGAGCATTTGGATGATCATGTTCTTGTGCGTGGGAATGAGTCA 461
Qy 444 AlaLeuTyrAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db 462 GCTCTTTGGGATTGCAACATCATGATGATGGGGAACATAGTAACTGCTACTCACCACAA 521
Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db 522 GATGCTGAGTGCACCTGCTCAGATGGATCCAACTTTGGAATGAGGCTGACCGCTGGAGG 581
Qy 483 SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyLeuTrpGlyThrValCysHis 502
```

Db 582 AATATGTTCTGGAAGAATAGAGATCAAAATCCAAAGGACGGTGGGGAACAGTGTGTGAT 641  
QY 503 AspArgTrpSerThrArgAsnAlaAlaValValCysGlnLeuGlyCysGlyLysPro 522  
Db 642 GATAACTCTCAACATAGATCATGTGCTGCTTGTAGACAACTGAATGTGGAAGTCT 701  
QY 523 MethHisValPheGlyMetThrTyPheLysGluAlaSerGlyProIleTrpLeuAsp 542  
Db 702 GTTCAGTTCTCTGGTTCATCTAATTTTGGAGAAGGCTCTGGACCACTCTGGTTGATGAT 761  
QY 543 ValSerCysIleGlyAsnGlnSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562  
Db 762 CTTATATGCAACGGAATAGTCAGCTCTCTGGAACCTGCAACATCAAGGATGGGAAAG 821  
QY 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582  
Db 822 CATAACTGTGATCATCTGAGGATGCTGGAGTATTGCTCAAGGAGCAGATCTGAGC 881  
QY 583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyPheGlnGly 602  
Db 882 CTGAGACTGCTAGATGGAGTCACTGAATGTTTCAGGAAGATTACAGTGAGATTCCAAAG 941  
QY 603 ArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCysSer 622  
Db 942 GAATGGGGACAATATGTCATGACGGCTGGGACAGTTACCATGCTGCTGGCATGCAAG 1001  
QY 623 GlnLeuAspCysProSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642  
Db 1002 CAACTGGGATGTCCTCACTGCCGTCACAGCCATTGGTGGAGTTAACGCCAGTAAGGATTT 1061  
QY 643 GlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662  
Db 1062 GGACATCTGGCTTGACAGCGTTTCTGCCAGGACATGAACCTGCTGCTGGCAATGT 1121  
QY 663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682  
Db 1122 AAACACCATGAATGGGAAGCAATATTGCAATCACAATGAAGATGCTGGCTGACATGT 1181  
QY 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerArgCysAlaGlyLys 702  
Db 1182 TCTGATGGATCAGATCTGAGCTTAAGACTTAGAGGTGGAGCGCCCTGCTGGGACA 1241  
QY 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722  
Db 1242 GTTGAGGTGGAGATTTCAGAGACTGTTAGGAAGGTGTGTACAGAGGCTGGGACATCAA 1301  
QY 723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742  
Db 1302 GAAGCTGATGTGGTTGCAAGGACCTGGGATGTGGATCTGCACCTCAAAACATCTTATCAA 1361  
QY 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762  
Db 1362 GTGTACTCCAAATCCAGGCACAAACACATGGCTGTTCTAAGTAGCTGTAAACGGAAT 1421  
QY 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782  
Db 1422 GAAACTTCTTTGGGACTGCAAGAACTGGCAATGGGTGGACTTACCTGTGATCACTAT 1481  
QY 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802  
Db 1482 GAAGAAGCCAAAATTAACCTGCTCAGCCACAGGGAACCCAGACTGTTGGAGGGACATT 1541  
QY 803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822  
Db 1542 CCTGTGTCGGAGCTGTTGAAGTGAAGCATGTTGGACACGCTGGGCGCTCCATCTGTGATTCG 1601  
QY 823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842  
Db 1602 GACTTCTCTGGAAGCTGCCAGCGTTCTATGCGAGGAATATACAGTGTGGCAGCTTGTGC 1661  
QY 843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862  
Db 1661

Db 1662 TCTATCTCTGGGGAGCTCACTTTGGAGAGGGAATGCACAGATCTGGGCTGAAGAAATTC 1721  
QY 863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882  
Db 1722 CAGTGTGAGGACATGATGATCCATCTTCTACCTGTGCCAGTAGCACCCTCCGCCAAGAGGA 1781  
QY 883 ThrCysIleHisSerArgGluValGlyValValCysSerArgTyThrAspValArgLeu 902  
Db 1782 ACITGTAGCCACACAGCAGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTG 1841  
QY 903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922  
Db 1842 GTGAATGGCAAGACCCGCTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGGTGCCTGGGGA 1901  
QY 923 SerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942  
Db 1902 TCCTCTGTAACTCTCACTGGACATAGAAGATCCCATGTTCTTCCACAGACCTTAAA 1961  
QY 943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyIleGlyGluArgSerValArgVal 962  
Db 1962 TGTGGAGTTGCCCTTCTACCCAGGAGGAGCAGCTTTTGGAAAAGGAAATGGTCAGATC 2021  
QY 963 TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThr 982  
Db 2022 TGGAGGCATATGTTTCACTGACCTGGGACTGAGCAGCAGCAGATGTCCTGTAACT 2081  
QY 983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002  
Db 2082 GCTCTAGTGCTTCAATATGCTTCTCAGAGCAAGTGGCTCTGTAAATCTGCTCAGGAAC 2141  
QY 1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyThrLeuSerAlaVal 1022  
Db 2142 CAGTCCCAACACTGCTCCTCGTCAATTCATCGTCTTGGGCCCAACAAGCCCTACCATT 2201  
QY 1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042  
Db 2202 CCAGAAGAAAGTGTGTGGCTGTGATAGAGAGTGTCAACTTCGCCCTGGTAAATGGAGGA 2261  
QY 1043 SerArgCysAlaGlyArgValGluIleTyHisAspGlyPheTrpGlyThrIleCysAsp 1062  
Db 2262 GTGCTGTGCTGGGAGATGAGATCTATCATGAGGCTCTCTGGGCAACCATCTGTGAT 2321  
QY 1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082  
Db 2322 GACAGCTGGGACCTGAGTGTGATGCCACCTGCTTGGCAGACAGCTGGCTGTGGAGAGGCC 2381  
QY 1083 PheAsnAlaThrValSerAlaHisPheGlyGluClySerGlyProIleTrpLeuAspAsp 1102  
Db 2382 ATTAATGCCACTGCTGCTCATTTTGGGGAAGGAACAGGGCCCATCTGGCTGGATGAG 2441  
QY 1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122  
Db 2442 ATGAATGCAATGGAAAAGAAATCCCGCATTTGGCAGTGGCCATTCACAGCGCTGGGGCAG 2501  
QY 1123 HisAsnCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142  
Db 2502 CAAAATGGCAGGCACAAAGGAGGATGCGGAGTTATCTGCTCAGAATTCATCTCTCAGA 2561  
QY 1143 LeuTyrsSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrsAsnGly 1162  
Db 2562 CTGACCACTGAAGCCACAGCAGAGGCGCTGTGACAGGGGCTCTGGAAGTGTTCACAAATGGA 2621  
QY 1163 ThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCysArg 1182  
Db 2622 GCTTGGGGCACTGTTGGCAAGAGTAGCATGCTGAAACCACTGTGGGTGGTGTGGCAGG 2681  
QY 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202  
Db 2682 CAGCTGGCTGTGCAGACAAAGGAAATCAACCTGCATCTTTAGACAGGCCATGTCC 2741  
QY 1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222  
Db 2742 ATTCCTCATGTGGTGGACAATGTTCACTGTCCAAAAGGACCTGCACAGCTGTGGCAGTGC 2801

```

Qy 1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242
Db 2802 CCATCATCTCCATGGAGAGAGACTGGCCAGCCCTCGGAGGAGACTGGATGATCATGT 2861
Qy 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
Db 2862 GACACACAGTAAGACTTCAGAGAGCCCACTTCCTGTTCTGGACGTGGGATCTGG 2921
Qy 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluVal 1282
Db 2922 CATGGAGTTCTCGGGGACAGTGTGATGACTCTTGGGACTTGGAGGATGCTCAGGTG 2981
Qy 1293 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGly 1302
Db 2982 GTGTGTCACCAACTGGCTGTGGTCCAGCTTTGAAAGCATTCACAAAGACGAGAGTTTGGT 3041
Qy 1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322
Db 3042 CAGGGAGTACGACCGCATATGCTCAATGAAGTGAAGTCAAGGGAATGAGTCTTCCTGTG 3101
Qy 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
Db 3102 TGGGATTCTCTGCCAGAGCTGGGGCCATAGTGTGAGTGGGCACAGGACGCTGCA 3161
Qy 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356
Db 3162 GTGAATTGCACAGATATTTTCAGTGCAGAAAAACCCACAAAAAGCCACACAGGCTGCTCA 3221
Qy 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376
Db 3222 TCCCGTCAGTCATCTTTATGTCAGTGGGATCCTTGGGGTGTCTCTTGGCCATTTTC 3281
Qy 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
Db 3282 GTCGATTTCTTCTGACTTAAAGCGAAGACAGACAGACAGCGG-----CTTGCA 3332
Qy 1395 ValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
Db 3333 GTTCTCTCAAGAGGAGAGAACTTAGTCCACCAAAATTCATACCGGGAGATGAATCTTTCG 3392
Qy 1415 LeuLysArgGluAsp 1419
Db 3393 CTGAATGACAGATGAT 3407

RESULT 7
HSM130AE HSM130AE 3802 bp mRNA linear PRI 22-NOV-1996
DEFINITION H.sapiens mRNA for M130 antigen extracellular variant.
ACCESSION Z22971
VERSION Z22971.1 GI:312147
KEYWORDS antigen; antigen M130.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department
of Pathology, University of Oxford, Level 1, Maternity Block, John
Radcliffe Hospital, Headington, Oxford, OX3 9DU, United Kingdom
REFERENCE
AUTHORS Law, S.K., Micklem, K.J., Shaw, J.M., Zhang, X.P., Dong, Y., Willis, A.C.
and Mason, D.Y.
TITLE A new macrophage differentiation antigen which is a member of the
scavenger receptor superfamily
JOURNAL Eur. J. Immunol. 23 (9), 2320-2325 (1993)
MEDLINE 93380506
PUBMED 8370408
FEATURES
Location/Qualifiers
1..3802
/organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/cell_type="Macrophage"
102..3551
/codon_start=1
/product="M130 antigen extracellular variant"
/protein_id="CAA80544.1"
/db_xref="GI:312148"
/db_xref="SPTREMBL:Q07901"
/translaton="MVLLESGSADFRHRFVNLSPTITVVLLLSACFVTSLSGGTOK
ELRLVDENKCSGRVEYKQVEWGTVCNNGSMSEAVSICNOLGCPATIAKAPGWNSS
AGSGRIWMDHVSCGRNBSALDCKHDGKHSNCTHOQDAGVTCSDGNSNLEMRITRG
NMCSGRIEIKFOGRWGTVCDDNFENIDHASVICROLECGSAVSFGSSGNFEGSGPIWF
DDLICNGESALNWKCHQKNGKNCDAEDAGVTCCKGADLSRLVDGVTECSGRLEV
RFQEWGTICDDGWDSDYDAAVACKQLGCPATVTAIGRNASKGFGHILDSVSCQGH
PAYQCKHHEWGHYCNHEDAGVTCSDGSDLELRRLGGSGRCAGTVEVEITRLGKV
CDRGWLTKEADVVCROLGCSALKTSQVYTKQATNTWLFSSCNGNETSLMDCKNM
QWGLTGDHYEAKITCSAHRPELVGGDIPKSGRVEVKGHDTWGSICDSDFSLAAS
VLCRELOCQTVVSTLGGARHGGEGQIWAEEFOCEGHESHLCLCPVAPRPGTCSHSR
DVGWVSSKTQKTSLIGSYVYKGTGLSHSCLFLKPCLLPGYTEIRLVNKGTPCEGRV
ELKTLGAWGSLCNSHWDIEAHVLCQQLKCGVALSTPGGARFGNGQIWRHMFCTG
TEHMGDCPTALGASCPSEQVAVSVCNQSOTLSSCNSSSLGTPRTPIPEESAVA
CIESGQLRVNGGRCAGRVEIYHEGWSGTICDDSDLSDAHVVYCRQLGCGEALNATG
SAHFEGCTPIWLDENKCGKESRWOCHSHGWGQONCRKEDAGKINPALSIDRAMSI
PMWVDVQCCKGPDYTLQCPSSPNEKRLASPSERTWITCDNKIRLOEGPTSCSRVEI
WHGWSGTVCDDSDWDLDAQVYCOQLGCGPALKAFKAEFGGQGTGPIWLNVECKNGV
SSLWDCPFARMHGHSECKHEDAAVNCDDISVQKTPKQKATGRSSRSQSFIAVGLGVV
LLAIFVALFELTKRRQRQLAVSSRGENLVHQIYQREMNSCLNADDLMLNMSGGHS
EPH"
sig_peptide 102..221
mat_peptide 222..3548
/product="M130 antigen extracellular variant"
/evidence=experimental
polyA_signal 3776..3781
polyA_signal 3780..3785
polyA_signal 3784..3789
polyA_site 3799
BASE COUNT 1009 a 751 c 1081 g 961 t
ORIGIN
Alignment Scores:
Pred No.: 2,41e-212 Length: 3802
Score: 3396.50 Matches: 597
Percent Similarity: 69.03% Conservative: 161
Best Local Similarity: 54.37% Mismatches: 295
Query Match: 41.74% Indels: 45
DB: 9 Gaps: 5
US-09-759-130B-381 (1-1453) x HSM130AE (1-3802)
Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db 222 GGAACAGACAAGGAGCTGAGGCTAGTGATGGTGAACAAAGTAGTCGCGGAGAGTGGA 281
Qy 384 ValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla 403
Db 282 GTGAAGTCCAGGAGGAGTGGGACGCTGTATAATGGCTGGAGCATGGAACGGTTC 341
Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423
Db 342 TCTGTGATTTGTAAACACCTGGATGTCCTCAACTGCTATCAAGCCCTGGATGGCTAAT 401
Qy 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db 402 TCCAGTGCAGGTTCTGGACGACATTTGGATGATCATGTTCTTCTGCTGGGAATGAGTCA 461
Qy 444 AlaLeuTrpAspCysThrTyrrAspClyLysAlaLysAlaArgThr---CysPheArgArgSer 462
Db 462 GCTCTTTGGGATTCGAAACATGATGGGAAAGACAGTACTGACTACTACCAACAA 521
Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuValGlyAlaHis 482
Db 522 GATGCTGAGTGACCTGCTCAGATGATCCCAATTTGGAAATGAGGCTGACCGCTGGGG 581

```

QY 483 SerProCysTyrGlyArgLeuGluValIleYstYrGlnGlyGluTrpGlyThrValCysHis 502  
DB 582 AATATGTGTTCTGGAAGAAATAGAGATCAAAATCCAAAGGCGGTGGGAACAGTGTGTAT 641  
QY 503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522  
DB 642 GATAACTGTCATCATGCTGATGCTGCTCATTTGTAGAACATTTGAATGTGAAGTGT 701  
QY 523 MethHisValPheGlyMetThrTyrPheIysGluAlaSerGlyProIleThrLeuAspAsp 542  
DB 702 GTCAAGTTTCTGTTGCTTAAATTTTGGAGAGGCTCTGGACCCTTGTGTTGATGAT 761  
QY 543 ValSerIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562  
DB 762 CTTATATGCAACGGAATAGTACAGTCTCTGGAACCTGCAACATCAAGGATGGGAAG 821  
QY 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582  
DB 822 CATAACTGCTGATCATGCTGAGGATGCTGAGTGAATTTGCTCAAAAGGAGCAGATCTGAGC 881  
QY 583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602  
DB 882 CTGAGACTGGTAGAGTACCTGAATGTTTCAGGAAGATTAGAAAGTGAGATCCCAAGGA 941  
QY 603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSer 622  
DB 942 GAATGGGGGACAATATGATGATGACGGCTGGGACATTTACGATGCTGCTGGCATGCAAG 1001  
QY 623 GlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642  
DB 1002 CAACCTGGGATGCTCCAACTGCGCTCACAGCCATTTGCTGAGTTAACGCCAGTAAAGGATTT 1061  
QY 643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662  
DB 1062 GGACACATCTGGCTTGACAGCTTCTTGCCAGGACATGAACCTGCTGCTGCAATGT 1121  
QY 663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682  
DB 1122 AAACACCATGAATGGGAAGACATATTATGCAATCAATGAAGATGCTGGCGTGACATGT 1181  
QY 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702  
DB 1182 TCTATGATGATGATCTGGAGCTTAAGACTTAGAGGTGGAGCAGCCGCTGTGCTGGACA 1241  
QY 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722  
DB 1242 GTTCAGGTGGAGATTCAGACTCTTAGGGAAGGTGTGTGACAGGCTGGGACTGAAA 1301  
QY 723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742  
DB 1302 GAAGCTGATGTGTTTGCAGGCAGCTGGGATGTGATCTGCACATCAAAACATCTATCAA 1361  
QY 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762  
DB 1362 GTGTATCCCAAAATCCAGGCCAACACATACATGCTTCTTAAGTAGTGTGAACGGAAT 1421  
QY 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782  
DB 1422 GAACTTCTCTTTGGGACTGCAAGACTGGCAATGGGGTGGGCTTACCTGTGATCACTAT 1481  
QY 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802  
DB 1482 GAAGAAGCAAAATTAACCTGCTAGCCACACAGGGAACCCAGACTGTTGGAGGGGACAT 1541  
QY 803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822  
DB 1542 CCCTGTTCTGACGCTGTGAAGTGAAGCATGGTGACACATGGGGCTCCATCTGTCAATCG 1601  
QY 823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842  
DB 1602 GACTTCTCTGGAAGCTGGCAGGCTTATGCAAGGGAATTTACAGTGTGGCAGCATGTTGTC 1661

QY 843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862  
DB 1662 TCTATCTCTGGGGGAGCTCACTTTGGAGAGGAAATGACAGATCTGGGCTGAAGAATTC 1721  
QY 863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882  
DB 1722 CAGTGTGAGGAGCATAGTCCCATCTTCACTCTGCCAGTAGCACCCCGCCCAAGAAG 1781  
QY 883 ThrCysIleHisSerArgGluValGlyValValCysSerArg 896  
DB 1782 ACTTGTAGCCACACAGGATGTTGGAGTAGTCTGCTCAAGTAGACCCCAAAAACATCT 1841  
QY 896 896  
DB 1842 TTAATTTGTTCTTATATCTGTGAAAGGACAGGGTTAGGGAGTCATAGCTGTCTTTTCTA 1901  
QY 897 897  
DB 1902 AAGCCCTGCTCTCTCCAGGATACAGAAATTCGTTGGTGAATGGCAAGACCCCGTGT 1961  
QY 910 AspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp 929  
DB 1962 GAGGCGAGAGTGGAGCTCAAAACGCTTGGTCCCTGGGATCCCTCTGTAACCTCTCACATGG 2021  
QY 930 AspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThr 949  
DB 2022 GACATGAAGATGCCATGTTCTTTGCCAGCAGCTTAAATCTGGAGTTGCCCTTTCTACC 2081  
QY 950 ThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCys 969  
DB 2082 CCAGGAGGAGCAGCTTTTGGAAAGAAATGTCAGATCTGGAGCATATGTTTCACTGC 2141  
QY 970 LeuGlyAsnGluSerLeuLeuAspAsnGlnMetThrValLeuGlyAlaProCys 989  
DB 2142 ACTGGACTGAGCAGCATGGAGATGTCCTGTAACCTGCTAGGTGCTTCAATATGT 2201  
QY 990 IleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGlnProLeuPhePro 1009  
DB 2202 CTTTCAGAGCAAGTGGCTCTGTAATCTGTCAGAAACACAGTCCCAACACATGTCCTCG 2261  
QY 1010 CysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIle 1029  
DB 2262 TGCAATTCATCGTCTTTGGGCCCAACAAGGCTTACCATTCCAGAAAGAGTCTGTGGCC 2321  
QY 1030 CysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgVal 1049  
DB 2322 TGCAATGAGAGTGTCACTTCGCTGTAATGGAGGAGGTGCTGCTGGGAGAGTA 2381  
QY 1050 GluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069  
DB 2382 GAGATCTATCATGAGGGCTCTGGGGCACCATCTGTGATGACAGCTGGGACCTGAGTGT 2441  
QY 1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaAlaPheAsnAlaThrValSerAla 1089  
DB 2442 GCCACAGTGGTTCAGACAGCTGGGCTGTGGAGAGGCCATTAATGCCACCTGGTCTCT 2501  
QY 1090 HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109  
DB 2502 CATTTGGGGAAGAACAGGCCCATCTGCTGGTGGATGAGATGAATGCAATGGCAAGAA 2561  
QY 1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129  
DB 2562 TCCCGCATTTGGCAGTGCCATTCACACGCTGGGGGACAGAAAATTCAGGACCAAGAG 2621  
QY 1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThr 1149  
DB 2622 GATCGGGAGTATCTGCTCAGAAATTCATGCTCTGAGACTGACACATGCAAGCCAGCAGA 2681  
QY 1150 GluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArg 1169  
DB 2682 GAGGCTGTGCGGCGCTCGAAGTTTTTTTACATGAGGCTTGGGCGACCTGTGGCAG 2741  
QY 1170 ArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189

Db 2742 AGTAGCATGCTGTAACCAACTGTGGGTGTGGTGTGACGACGCTGGGTGTGACACAAA 2801  
Qy 1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAsp 1209  
Db 2802 GGGAAATCAACCCCTGATCTTTAGACAGGCCATGTCCATTCCTCATGTGGTGGACAA 2861  
Qy 1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229  
Db 2862 GTTCAGTGTCCAAAGGACCTGACACGCTGTGGCAGTGGCCATCATCTCCATGGGAGA 2921  
Qy 1230 ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg 1249  
Db 2922 AGACTGGCCAGCCCTCGGAGAGACCTGGATCATGTGACACAAAGATAAGACTTCAG 2981  
Qy 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269  
Db 2982 GAAGAGCCCACTTCCTGTCTGGAGCTGGAGATCTGGCAGGAGTTCCTGGGGGACA 3041  
Qy 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys 1289  
Db 3042 GNGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTGGTGTGTCACAACTTGGCTGT 3101  
Qy 1290 GlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309  
Db 3102 GGTCCAGCTTTGAAAGCATTCACAAAGACAGAGTTGTGTCAGGGGACTGGACCGATATGG 3161  
Qy 1310 LeuAspAspMetArgCysLysGlyAsnGlnLeuTrpPheLeuTrpAspCysHisAlaLysPro 1329  
Db 3162 CTCATGAGTGAAGTGCACAAAGGAATGAGTCTTCTGTGGGATGTCCTGCCAGAGCC 3221  
Qy 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349  
Db 3222 TGGGGCCATAGTGTGGGCGACAAAGAGAGCGTGCAGTGAATGTCACAGATATTTCA 3281  
Qy 1350 LeuLys-----SerLeuAsnAlaSerSerGlyHisLeuAlaLeuIle 1363  
Db 3282 GTGCAGAAACCCCAAAAGCCACACAGGCTCATCCGTCAGTCATCCTTTAT 3341  
Qy 1364 LeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIle-----LeuPheLeuThr 1381  
Db 3342 GCAGTCGGGATCCTTGGGGTGTCTGTGGCCATTTTCGTCGCATATTCTCTTGACT 3401  
Qy 1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgGly 1401  
Db 3402 AAAAGCGAAGACAGACAGACGCG-----CTTGCAAGTTCCTCAAGAGGAGAGA 3452  
Qy 1402 SerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAsp 1419  
Db 3453 TTAGTCCACCAATTCATACCGGAGATGAATCTTGCTGATGACAGATGAT 3506  
RESULT 8  
LOCUS AX281768 3707 bp DNA linear PAT 02-NOV-2001  
DEFINITION Sequence 177 from Patent WO017399.  
ACCESSION AX281768  
VERSION AX281768.1 GI:16609019  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1  
Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,  
Mikita,T. and Tai,J.  
Genes expressed in foam cell differentiation  
Patent: WO 0177389-A 177 18-OCT-2001;  
Incyte Genomics, Inc. (US)  
Location/Qualifiers  
1..3707  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 236574.12"

BASE COUNT 984 a 731 c 1062 g 930 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4.38e-211 Length: 3707  
Score: 3377.00 Matches: 598  
Percent Similarity: 71.00% Conservatives: 161  
Best Local Similarity: 55.94% Mismatches: 294  
Query Match: 41.50% Indels: 16  
DB: 6 Gaps: 4  
US-09-759-130B-381 (1-1453) x AX281768 (1-3707)  
Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSer-GlyArgValGl 383  
Db 222 GGAACAGACAAGGAGCTGAGGCTAGTGTGTAACAAACAGCTAGCGGGGAGAGTGA 281  
Qy 383 uValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAl 403  
Db 282 AGTGAAGCTCCAGGAGGAGTGGGAAACGCTGTGTATATATGGCTGGAGCATGGAACCGGT 341  
Qy 403 aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423  
Db 342 CTCGTGTATGTATACACAGCTGGGATGTCCAACTCTATCAAAAGCCCTGGATGGCTAA 401  
Qy 423 sProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe 443  
Db 402 TTCAGTGCAGGCTTCTGGACGCAATTTGGATGGATCATGTTCTTGTGCTGGGAATGAGTC 461  
Qy 443 rAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThr---CysPheArgArgSe 462  
Db 462 AGCTCTTTGGGATTCACAAACATGATGGATGGGAAAGCATAGTAATCTACTACCAACA 521  
Qy 462 rAspAlaGlyValIleCysSerAspLysAlaAspLeuArgLeuValGlyAlaHi 482  
Db 522 AGATCTGGAGTGACCTGCTCAGATGGATCCAAATTTGNAATGAGCGTGACCGTGGAGG 581  
Qy 482 sSerProCysTyArgLeuGluValLysTyArgGlnGlyGluTrpGlyThrValCysHi 502  
Db 582 GAATATGTCTTGTGAAGAATAGAGATCAAAATTCGAAGGACGCTGGGGAACAGTGTGA 641  
Qy 502 sAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGlyLysPr 522  
Db 642 TGATAACTCAACATAGATGATGCTCATTTGTAGACAACTGAATGATGGAGAGTGC 701  
Qy 522 oMetHisValPheGlyMetThrTyPheLysGluAlaSerGlyProIleTrpLeuAspAs 542  
Db 702 TGTCAGTTTCTGCTTCATCTAATTTTGAGAGAGCTCTGGACCAATCTGCTTTGATCA 761  
Qy 542 pValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLy 562  
Db 762 TCTTATATCCAAACGGAATGATGATGCTGAGGATGCTGAGTGTGCTCAAGGGGAGAGTCA 821  
Qy 562 shiAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGl 582  
Db 822 GCATAACTGTGATCATGCTGAGGATGCTGAGTGTGCTCAAGGGGAGAGATCTGAG 881  
Qy 582 yLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyPheGlnGl 602  
Db 882 CCGTAGACTGGTGTAGATGAGTCACTGAATTTTCAGGAAGATTAGAAGTGGAGATTCAGG 941  
Qy 602 yArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCysSe 622  
Db 942 AGAATGGGGGACAAATGTGATGACGGCTGGGACAGTTACGATGCTGCTGGCATGCA 1001  
Qy 622 rGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTy 642  
Db 1002 GCAACTGGGATGTCCAACCTGCGCTCACAGCCATTTGGTCAGTTAACGCCAGTAAGGAT 1061  
Qy 642 rGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCy 662  
Db 1062 TGGACACATCTGGCTTGACAGCGTTTCTTGCCAGGAGACATGAACCTGCTGTGCTGCAATG 1121

Qy 662 sArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682  
Db 1122 TAAACACCAGTAATGGGAAAGCATTATTGCAATCAATGAAGATGCTGGCGTGACATG 1181  
Qy 682 sSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerArgCysAlaGlyLy 702  
Db 1182 TTCTGATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGCGCCGTGCTGGGAC 1241  
Qy 702 sValGluValAsnValGlnAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAs 722  
Db 1242 AGTTGAGCTGGAGATTCAGAGACTGTAGGGAAGGTGTGTGCAGAGAGCTGGGGAGTAA 1301  
Qy 722 nIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArg 742  
Db 1302 AGAAGCTGATGTGTTTGCAGGACGTGGATGTGGATCTGCACCTCAAAACATCTTATCA 1361  
Qy 742 uproHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyG 762  
Db 1362 AGTGATCTCCAAATCCAGGCCAACAAACACATGGCTGTTTCTAAGTAGCTGTAAACGGAAA 1421  
Qy 762 yGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAs 782  
Db 1422 TGAACACTCTCTTGGGACGTGCAAGACTGGCAATGGGTGGAGTACCTGTGTATCACTA 1481  
Qy 782 nMetGluAlaSerLeuLeuCysSer-AlaHisArg-GlnProArgLeuValGlyAlaAsp 801  
Db 1482 TGAAGAAGCCAAATACCTGCTCAGGCCACACAGGGGAAACCCAGACTGGTTGGAGGGAC 1541  
Qy 802 MetProCysSerGlyArgValGluValValHisAlaAspThrTrpArgSerValCysAsp 821  
Db 1542 ATTCCTCTGTTGGACGTGTGAAGTGAAGCATGGTGACACGTGGGGCTCCATCTGTGAT 1601  
Qy 822 SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnGlyValAspAla 841  
Db 1602 TCACACTCTCTCTGGAGGTGCCAGGTCTATGCAGGGAATACAGTGTGGCACAGTT 1661  
Qy 842 IleSerLeuSerValGlyAspHisPheClyLysGlyAsnGlyLeuThrTrpAlaGluLys 861  
Db 1662 GTCTCTATCTCTGGGGGAGCTCACTTTGGAGAGGGAAATGACAGATCTGGGCTGAAGAA 1721  
Qy 862 PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu 881  
Db 1722 TTCAGTGTGAGGACATGATGCCATCTTCACTCTGCCAGTAGCACCCGCCCCAGAA 1781  
Qy 882 AspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspValArg 901  
Db 1782 GGAATTTGTAGCCACACAGCGGATGTGGAGTAGTCTGCTCAAGATACACAGAAATTCG 1841  
Qy 902 LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp 921  
Db 1842 TTGTGAATGGCAAGACCCCGTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGGTGCCTGG 1901  
Qy 922 GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu 941  
Db 1902 GGATCCCTCTGTAACTCTCACTGGGACATAGAAGATGCCATGCCATGTCTTGGCAGAGCT 1961  
Qy 942 SerCysGlyThrAlaLeuSerThrGlyGlyLysTrpIleGlyGluArgSerValArg 961  
Db 1962 AAATGTGGAGTGGCCCTTCTACCCAGGAGGACAGCTTTTGGAAAGAAATGGTCA 2021  
Qy 962 ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet 981  
Db 2022 ATCTGGAGGATATGTTTCACTGACGTGGGACTGAGCAGCAGCATGGGAGATTGTCCTGTA 2081  
Qy 982 ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly 1001  
Db 2082 ACTGCTCTAGTGTCTTATTATGTCCTTCAGCAAGTGGGCTCTGTAATCTGTCTCAGGA 2141  
Qy 1002 SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTrpLeuSerAla 1021  
Db 2142 AACAGTCCCAACACACTGCTCTGTGCAATTCATCTGCTTTGGGCCCAACAGCCCTACC 2201  
Qy 1022 ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly 1041

Db 2202 ATTCACCAAGAAAGTGTGTGGCCTGCATAGAGAGTGGTCAACTTCGCTGTAATGGA 2261  
Qy 1042 AspSerArgCysAlaGlyArgValGluIleTrpHisAspGlyPheTrpGlyThrIleCys 1061  
Db 2262 GGAGGTGCTGTGTGGGAGTAGAGATCTATCATGAGGCTCTGGGGACCATCTGT 2321  
Qy 1062 AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal 1081  
Db 2322 GATGACAGCTGGGACCTGAGTATGCCACGTGGTTGGACAGAGCTGGGCTGTGGAGAG 2381  
Qy 1082 AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp 1101  
Db 2382 GCCATTAAATGCCACTGCTTCTGCTCATTTTGGGAAGAACAGGGCCCATCTGGCTGAT 2441  
Qy 1102 AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly 1121  
Db 2442 GAGTGAATGAATGGAAGAAATCCCGCATTTTGGCAGTCCCATCACAGCTGGGGG 2501  
Qy 1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141  
Db 2502 CAGCAAAATTCAGGCACAGGAGGATGCCGGAGCTTATCTGCTCAGAAATTCATGCTCTG 2561  
Qy 1142 ArgLeuTrpSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTrpAsn 1161  
Db 2562 AGACTGACCACTGAAGCCAGCAGAGAGGCGCTGTGCAGGCGCTCTGGAAGTCTTTTACAAT 2621  
Qy 1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCys 1181  
Db 2622 GGAGCTTGGGGCACTGTGGCAAGAGTAGCATGCTGTAACACCACTGTGGGTGGTGTGTC 2681  
Qy 1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGly 1201  
Db 2682 AGGCAGCTGGCTGTGCAGCAAAAGGGAATCAACCTGTCATCTTTAGACAGCCCATG 2741  
Qy 1202 SerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp-G 1221  
Db 2742 TCCATTCCCATGTGGTGGCAATGTTCAAGTGTCCAAAGGACCTGCACACGCTGTGGCA 2801  
Qy 1221 nCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleTh 1241  
Db 2802 GTGCCCATCATCTCCATGGGAGAGAGACTGCCAGCCCTCGGAGGAGACCTGGATCAC 2861  
Qy 1241 rCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluI 1261  
Db 2862 ATGTGACAAACAAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGCTGGAGAT 2921  
Qy 1261 eTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaG 1281  
Db 2922 CTGCATGGAGTCTCTGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTPCA 2981  
Qy 1281 uValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPh 1301  
Db 2982 GGTGGTGTCTCAACAACCTGGCTGTGGTCCAGCTTTGAAAGCATTCACAAAGACGAGATT 3041  
Qy 1301 eGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPh 1321  
Db 3042 TGGTCAGGGGACTGGACGATATGGCTCAATGAAGTGAAGTGAAGGAAGGAATGAGTCTTC 3101  
Qy 1321 eLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAl 1341  
Db 3102 CTCTGGGATGTCTGCCACAGCTGGGGCCATAGTGAAGTGTGGGCACAAGGAAGACGC 3161  
Qy 1341 aGlyValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAl 1355  
Db 3162 TGCAGTGAATTCAGATATTTTCAGTGCAGAAAAACCCCAAAAGGACCAACAGGTGCG 3221  
Qy 1355 aSerSerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLe 1375  
Db 3222 CTCATCCGTCAGTCATCTTTATTTGAGTGGGATCTCTGGGGTGTCTTGTGGCCAT 3281  
Qy 1375 uPheIle-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLe 1393  
Db 1393

```

Db 3282 TTTCGTGCGCATTTCTTCTTCTGCTAAAGCGAGACAGACAGACGCGG-----CT 3332
QY 1393 uArgValSerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluTh 1413
Db 3333 TGCAGTTTCTCAAGAGAGAGAACTTAGTCCACCACAAATTCATACCGGAGATGAATC 3392
QY 1413 rCysLeuLysArgGluAsp 1419
Db 3393 TTGCCTGAATGCAGATGAT 3411

RESULT 9
LOCUS AF274883 4405 bp mRNA linear ROD 10-MAY-2001
DEFINITION Mus musculus macrophage hemoglobin scavenger receptor CD163
ACCESSION AF274883
VERSION AF274883.1 GI:13236130
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4405)
AUTHORS Schaefer, D.J., Boret, F.S., Hongegger, A., Poehler, D., Linnseheid, P.,
Staege, H., Muller, C., Schoedon, G. and Schaffner, A.
TITLE Molecular cloning and characterization of the mouse CD163
homologue, a highly glucocorticoid-inducible member of the
scavenger receptor cysteine-rich family
JOURNAL Immunogenetics 53 (2), 170-177 (2001)
MEDLINE 21243729
PUBMED 11345593
REFERENCE 2 (bases 1 to 4405)
AUTHORS Schaefer, D.J., Boret, F.S. and Schaffner, A.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2000) Departement of Medicine, Medizinische
Klinik B, University Hospital Zurich, Ramistrasse 100, Zurich, ZH
8091, Switzerland
FEATURES
location/Qualifiers
1..4405
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="6"
/map="15.43 centirays distal from the D6Mit300"
85..3450
/feature="FAT; member of the scavenger receptor cysteine rich
type B superfamily; transmembrane protein type I"
/codon_start=1
/product="macrophage hemoglobin scavenger receptor CD163
precursor"
/protein_id="RAK16065.1"
/db_xref="GI:13236131"
/translation="MGHMRVLLGGAGSGPGKRFVHLGFFVAVVSSLLSASAVTNAPG
EMKRLRAGGNNCSGRVELEKIDKWTGVCNGSMNVEVSVVQVQCGPTSIKALGW
ANSAGSYIWMKVSCGTGNSALMDKHDGKHNCTHEKDGATVQCGSLNLEMLV
NSAGHCLGRVEIKFOGKWTGVDNDFSKDASVTCOLGCGSALSFSGSAKLGAGC
PIWDLACNENESALWCKHKGKHCNDHAEDVGVICLEGADLSRLVDCVRCSG
RLEVFQGEWGTVDNDLDRDASVVCQGLGCTAISAGIRVNASESGQIWLNDISC
EGEATLMECKHQBWKHCHREDAGVTCSDGADLELRVGGSGRCAIGIVEIQKL
TKMCSGTLTADADVVCQLCGSALQTAIYKSTGATNTWLPFGSCNGNETTFWQ
CKNWGGLCNDNFEAKVTCSGHREPLVGGETPCSGRVEKHWGVMSYCDFDLSL
EASVVCVLCQGTWVSTLGAHFGEGSGQWGEFQCGSDSHLSLCSVAPPLDRTC
THSRVSVVCYRIDIRLAGVSSGSGRVELKTIGANGPLCSHMDMEDAHVLCQKL
CGVAOSIPGAHFGAGGVWSHMFCTGTEEHIGDCLMTALGAPTCEGQVAVCS
GNQSTLLPCSSLSPVOTTSIPKESEVPCIASGQLRVLGGRCARVEVCHESG
GTVCDDNMDTDANYVCQLDGVAINATGSAIFEGAGAIWLDVICTGKSHIWCQ
HSHGWRCKRHKEDAGVTCSEFMSRLTNAHKENCTGRLEVFVNGTWGSGISNMS
PTTVGVCRLQGLCADNVTKIPSDKTPSRMVDVRCQPKGVDTLMOCPSPKMQRO
ASPQSQSEWIIICDNKIRLQEGHTDCSGRVEIWHKGSWGTVCDDSHDLNDKAVKQGL
CQAVKALRAEFGPTGIWLNKIKRGNESLWDCPKAPKWSHSDCKEDASIQCL
VOYQEMDSKADLDLLKSSGVQIRHTEKENDL"
85..198
sig_peptide
misc_feature 3220..3288

```

```

BASE COUNT 1174 a 884 c 1246 g 1101 t
ORIGIN
Alignment Scores: 3.85e-204 Length: 4405
Pred. No.: 3272.00 Matches: 582
Score: 68.11% Conservative: 174
Percent Similarity: 52.43% Mismatches: 332
Best Local Similarity: 40.21% Indels: 22
Query Match: 10 Gaps: 9
DB: 10

US-09-759-130B-381 (1-1453) x AF274883 (1-4405)
QY 320 GlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrp 339
Db 91 GGACACAGAATGGTCTTCTTGGAGGTGCTGCATCTCCTGGTGTGTAAGAGTTGCTGCT 147
QY 340 AspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerVal 359
Db 148 -----CAPCTAGGTCTTCTTCTTGGTGTGCTGCTCCTCTCTCAGTCCCTCTGCT 198
QY 360 IleCysSerAspGlyAlaAspLeu-----GluLeuArgLeuAlaAspGlySerAsnAsn 377
Db 199 GTCACTAAGCTCTGGAGAAATGAGAGAGAACTGAGCTGCGGGTGTGTAAGAACAC 258
QY 378 CysSerGlyArgValGluValArgIleHisGluGlnTrpThrIleCysAspGlnAsn 397
Db 259 TGTACTGGGAGAGTGAATTAAGATCCATGACAAAGTGGGGCACAGTGTGCAGTAACGGC 318
QY 398 TrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPhe 417
Db 319 TGGACATGAATGAAGTCTCCGTGTTTCCAGCAGCTGGGATGCCCACTTCTATTAAA 378
QY 418 GlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSer 437
Db 379 GCCTTGGATGGCTTAACCTCCAGCGCGCTCTGGATATATCTGGATGACAAAGTTTCT 438
QY 438 CysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyArgPlyLysAlaLysArgThr 457
Db 439 TGTACAGGGAATGAGTCAGCTCTTGGGACTGCAACATGATGGTGGGAAAGCATAAC 498
QY 458 CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArg 477
Db 499 TGTACCCATGAAGAAGATGCTGAGTGACCTCTCAGATGGATCTAATTTGGAGATGAGA 558
QY 478 LeuVal---GlyAlaHisSerProCysTyArgGlyArgLeuGluValLysTyArgGlyGlu 496
Db 559 CTGGTGAACAGTGGCGGCCACCGATGCTTAGCAAGAGTAGAATAAAGTTCCAGGAAAG 618
QY 497 TrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGln 516
Db 619 TGGGGGACGCTGTGACGACACAACTTCAGAAAGATGACCTCTCTGTGATTTGTAACAG 678
QY 517 LeuGlyCysGlyLysProMethHisValPheGlyMetThrTyArgPheLysGluAlaSerGly 536
Db 679 CTGGATGGAGAGTCCCAATTAGTTTCTTGGCTCAGCTAAATTTGGGACTGGTCTGGA 738
QY 537 ProIleTrpLeuAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGlu 556
Db 739 CCAATCTGGCTGATGACCTGCGCATGCAATGGAATGAGTCACTCTCTGGGACTGCAAA 798
QY 557 HisSerGlyTrpGlyLysHisAsnCysValHisArgLysAspValIleValThrCysSer 576
Db 799 CACCGGGGATGGGGAAGCATAACTGTGACCATGCTGAGGATGTGCGGTGTGATTTGCTTA 858
QY 577 GlyAspAlaThrTrpGlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeu 596
Db 859 GAGGAGACAGATCTGAGCTGAGATGTAGTGGATGGAGTCTCCAGATGTTTCAGAGATTG 918
QY 597 GluValTyPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerIleSala 616
Db 919 GAAGTGAGATTCAGAGGAATGGGGACCGGTGTGTGATGATGATAACTGGGATCTCCGGAT 978

```

QY	617	AlaAlaValValCysSerGlnLeuAspCysProSerSerIlelleGlyMetGlyLeuGly	636
DB	979	GCTTCTGTGGTGTGCAAGCAACTGGGATGTCCAACTGCCATCAGTCGATGGTGTGCGATT	1038
QY	637	AsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGlu	656
DB	1039	AATGCCAGTCAGGGATCTGGACAGATTGGCTTGACAACAATTCATCGGAAGACATGAG	1098
QY	657	SerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGlu	676
DB	1099	GCAACTCTTTGGAGTGTAAACACCAAGAGTGGGAAAGCATTTACTGTCATCATAGAA	1158
QY	677	AspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySer	696
DB	1159	GACGCTGGTGTGACATGTTCTGATGGACACATCTGGAACATTAGACTGTAGTGGAGGC	1218
QY	697	SerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAla	716
DB	1219	AGTCGCTGTGTCGCATGTGGAGGTGGAGATTCAAGAGCTGACTGGGAAGATGTGTAGC	1278
QY	717	AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAla	736
DB	1279	CGAGGCTGGACACTGGCAGATCGGATGTGTTTCACACAGCTTGCATGTGGATCTGCG	1338
QY	737	IleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsn	756
DB	1339	CTTCAAAACCCAGGCTAAGATCTACTCTAAAACCTGGGGCAACAATACGTGGCTCTTTCCT	1398
QY	757	SerGlyCysThrGlyGlyGlyAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGln	776
DB	1399	GGATCTTGTAAGAAATGAACACTACTTTTTGGCAATGCAAAACTGGCAGTGGGCGGC	1458
QY	777	ThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArg	796
DB	1459	CTTTCCTGTGATAAATTCGAAGAAGCCAAAGTTACTCTCAGGCCACAGGGAACCCAGA	1518
QY	797	LeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrp	816
DB	1519	CTGGTTGGAGGAGAATCCATCGTCTGGTGTGGGAAGTGAACACGGAGACGTGTGG	1578
QY	817	ArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeu	836
DB	1579	GGCTCCGCTGTGATTGTGACTTGTCTCTGGAGCTGCCAGTGTGGTGTGCAGGGAATTA	1638
QY	837	AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeu	856
DB	1639	CAATGTGGAAACAGTCGTCTCTATCTTAGGGGGGCACATTTTGGGAAGGAAGTGGACAG	1698
QY	857	ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIle	876
DB	1699	ATCTGGGGTGAAGAATTCCAGTGTAGTGGGATGAGTCCCATCTTTCACATGCTCAGTG	1758
QY	877	ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg	896
DB	1759	GCGCCCCCGCTAGACAGAACTTTGACCACACAGCGGATGTCCGCTAGTCTGCTCACGA	1818
QY	897	TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn	916
DB	1819	TACATAGATATTCGTCTGGCAGCGCGAGTCCCTCGTGTAGGGAAGAGTGGAGCTCAAG	1878
QY	917	ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal	936
DB	1879	ACACTCGGAGCCCTGGGGTCCCTCTGTGCAGTGTCTCATTTGGGACATGGAAGATGCTCATGTC	1938
QY	937	LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyLysTrpIleGly	956
DB	1939	TTATGTGCAGCAGCTGAAGTGTGGGTTCGCCAAATCTATTCCAGAAGGAGCACATTTTGGG	1998
QY	957	GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu	976
DB	1999	AAAGAGAGCTGGTCAGCTGTGGAGTCCACATGTTTCCACTGCATCGAAGTGGAGCAATATA	2058

Qy	977	AspAsnCysGlnmethrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer	996
Db	2059	GGAGATTGGCTCATGACTGCTCTGGGTGGCGAGTGTCCGAAGCAGAGTGGCGCTCT	2118
Qy	997	ValIleCysThrGlySerLeuThrGlnProLeuPheProCys-----LeuAlaAsnVal	1014
Db	2119	GTCACTCTGCAGAAACCAATCCACAGACATATTGGCATCTGATTGTCATGTCTCCAGTC	2178
Qy	1015	SerAspProThrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLys	1034
Db	2179	CAACAACA---AGCTCTACAATTCCAAAGGAGAGTGAAGTTCCCTGCAATGCAAGTGGC	2235
Qy	1035	ArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleThrHisAsp	1054
Db	2236	CAGCTTCGCTTGGTAGGTGGAGTGGTGGCTGGCTGGAGAGTGGAGGTCTGCCACGAG	2295
Qy	1055	GlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCys	1074
Db	2296	GGCTCTTGGGGCACCCTCTGTGATGACAATTGGGNATATGACTGATGCCAATGTGGTGTGC	2355
Qy	1075	GlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGly	1094
Db	2356	AAGCAGCTGGACTGTGGCTGGCAATTAAAGCCACTGGCTCTGCTACTTCCGGGAAGGA	2415
Qy	1095	SerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGln	1114
Db	2416	GCAGGAGCTATCTGGCTAGACGAAGTCACTGCACTGGGAAGAGTCTCATATATTGGCAG	2475
Qy	1115	CysProSerArgGlyTrpGlyGlnHisAspCysAtcGHisLysGluAspAlaGlyValIle	1134
Db	2476	TGCCATTCACTGGCTGGGAGGCCATACTCCAGGCACAAGAAGATGCAAGTGTATC	2535
Qy	1135	CysSerGluPheThrAlaLeuArgLeuTrpSerGluThrGluThrGluSerCysAlaGly	1154
Db	2536	TGCTCCGAGTTCACTGCTCTGAGGCTGACCAACGAAGCCACAAAGAAACACTGCACAGGT	2595
Qy	1155	ArgLeuGluValPheTrpAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThr	1174
Db	2596	CGCCTTCAAGTGTTTACAAATGGTACATGGGCACATATTGGCAGTAGCAATATATGCCCA	2655
Qy	1175	AlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu	1194
Db	2656	ACCAGCTGGGGTGGTGGCTGAGCTGGGCTGTGCAGACACAGGCGAGTGTGAACACC	2715
Qy	1195	AlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLys	1214
Db	2716	ATACCTTCACACAGACACCATCCAGGCCCATGTGGGTAGATCGTGTCCAGTGTCCAAA	2775
Qy	1215	ThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro	1234
Db	2776	GGAGTTGCACATTTGTGGCAGTGCCTCGTCACCTTGGAAACAGACAGAGCCAGCCCC	2835
Qy	1235	---AlaGluIuThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThr	1253
Db	2836	TCTCCAGGAGTCCGGATCATCTGTGCACAAACAAATTAAGACTCCAGGAAGGCGATACA	2895
Qy	1254	GluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAsp	1273
Db	2896	GACTGTCTGCAGCTGGGAGATCTGGCACAAAGCTTCTGGGGAACAGTGTGTGATGAC	2955
Qy	1274	SerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeu	1293
Db	2956	TCCTGGGATCTTAATGCTAAGGTGTGTATGTAAGCAGATTGGGCTGTGGCCAAGCTGTG	3015
Qy	1294	AlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspMet	1313
Db	3016	AAGGCATATAAAGAAGCAGCATTTGGTCAGGAATCTGGGCCCATATGGCTCAATGAAATT	3075
Qy	1314	ArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSer	1333
Db	3076	AAGTGTAGAGGAATGAGTCTTCCCTGTGGGATTTCTCTCCCAACACCGTGGATGCACAGC	3135
Qy	1334	AspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeu	1353

```

||||| 3136 GACTGTGGCACAAGAGAGTCTCCATCCAGTGCCTCCCAAAATGACTTTCAGAAATCA 3195
QY 1354 AsnAlaSerSerGlyHisLeuAlaLeuLeuSerSerIlePheGlyLeuLeuLeuLeu 1373
Db 3196 CATCATGGCAGAGTACACCCACCTCAGCGGACCTCTGGTTGGAGCCATCTATTG 3255
QY 1374 ValLeuPheIleLeuPheLeuThrTrpCys-----ArgValGlnLysGlnLys 1389
Db 3256 GTCCCTCCTATTGCTCTCCTGCTGGACTCTGAAGCGAGCAGGATTCAGCGA----- 3309
QY 1390 HisLeuProLeuArgValSerThrArgArgGlySerLeuGluLeuLeuLeuLeuLeu 1409
Db 3310 -----CTTACAGTTCTCTCAAGAGGAGAGGCTTCTTATCATCATCAAGTTCAGTACCAA 3360
QY 1410 GluMetGluThrCysLeuLysArgGluAsp 1419
Db 3361 GAGATGGATTC-----AAGCGGATGAT 3384

RESULT 10
LOCUS BBWC11MR 4475 bp mRNA linear NAM 24-MAY-1994
DEFINITION B.bovis WC1.1 mRNA.
ACCESSION X63723
VERSION X63723.1 GI:13
KEYWORDS WC1.1 protein.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 4475)
Wijnngaard,P.L., MacHugh,N.D., Metzelaar,M.J., Romberg,S.,
Bensaïd,A., Pepin,L., Davis,W.C. and Clevers,H.C.
Members of the novel WC1 gene family are differentially expressed
on subsets of bovine CD4-CD8- gamma delta T lymphocytes
J. Immunol. 152 (7), 3476-3482 (1994)
94194107
7511649
2 (bases 1 to 4475)
Wijnngaard,P.L.J.
Direct Submission
Submitted (15-JAN-1992) P.L.J. Wijnngaard, Dept of Immunology,
University Hospital, P O Box 85500, 3508 GA Utrecht, THE
NETHERLANDS
FEATURES
Source
1..4475
/organism="Bos taurus"
/strain="indicus"
/isolate="D409"
/db_xref="taxon:9913"
/clone="pBOWC1.1"
/cell_line="D409/N2"
/tissue_type="blood"
/clone_lib="D409/N2.H1"
/dev_stage="gamma/delta T-cell"
22..4332
/gene="WC1.1 subcluster gene"
22..4332
/gene="WC1.1 subcluster gene"
/codon_start=1
/product="BowC1.1"
/protein_id="CAA45255.1"
/db_xref="GI:14"
/translation="WALGRHLSRLGLCVLLITMTVGGQALRLKDGVRCEGRVEVK
HOGEWGTVDGYRWTLKDSVVCVRLGCGCAAGFGPGGLGFWLWLLYTSCEGTES
TVSDCEHSNIKIDRNDGNHGDAGVCSGVRLAGGDPCCSGRVEVHSGEAWIPVSD
GNFTLATQIAIEAGCGKAVSLVGHFRESSAQVWAEFECEGPELWVCPRVPC
PGCTCHHSAGQVSAYSSEVRLMTNGSQCEGVEMNISGWORALCALSHWSLANAY
ICRQLCGGAVLTPGPHLVEGDQILTAHFHCSGAESPLWSCPYTALGGDPDCSHGT
ASVICGNIQIOLPCNDVSQPTGSAASEDSAPYCSDSROLRLVDGGGPGCARGVEIL
DQSGNGTICDDGWDLDARVVCRLGCCGALNATGSAHFGAGSGFWLIDLNLNCTGKES

```

polyA\_site 4453  
BASE COUNT 933 a 1113 c 1442 g 987 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9 9e-192 Length: 4475  
Score: 2082.00 Matches: 613  
Percent Similarity: 55.85% Conservative: 208  
Best Local Similarity: 41.70% Mismatches: 495  
Query Match: 37.87% Indels: 154  
DB: 4 Gaps: 23

US-09-759-130B-381 (1-1453) x BBWC11MR (1-4475)

```

QY 28 CysIleLeuLeuLeuLeuSerCysPheLeuIleSerSerPheAsnGlyThrAspLeuGlu 47
Db 58 TGTGTCCTCCTCCTCGGC-----ACCATGGTGGTGGTCAAGCTCTGGAG 102
QY 48 LeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluValLysPheGlnGly 67
Db 103 CTGAGTGAAGGTGGAGTCCATCGCTGAGGGGAGAGTGAAGTGAAGCACCAGGA 162
QY 68 GlnTrpGlyThrValCysAspGlyTrpAsnThrThrAlaSerThrValValCysLys 87
Db 163 GAATGGGCGACAGTGGATGTTACAGGTGGACATTAAGAGTGCATCTCTAGTGTGCAGA 222
QY 88 GlnLeuGlyCysProPheSerPheAlaMet-----PheArgPheGlyGlnAlaVal 104
Db 223 CAGCTGGGTGTGGAGCTGCCATTTGCTTCTGGAGGGGCTTATTTTGGCGCAGGACTT 282
QY 105 ThrArgHisGlyLysIleTrpLeuAspValSerCysTyrGlyAsnGlnSerAlaLeu 124
Db 283 -----GGCCCCATTTGGCTTTTGTATCTTCTGATGAAGGACAGAGTCACTGTC 333
QY 125 TrpGluCysGlnHis-----ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu 141
Db 334 AGTGACTGTGACATCTTAATAAAGACTATCGTAATGATGGCTAATAATCATGTCGG 393
QY 142 AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn 161
Db 394 GATGCTGGAGTAGTCTGCTCAGGA-----TTTGTGCTGCTGCTGGAGGGAT 441
QY 162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181
Db 442 GGACCTCTGCTCAGGGCGAGTAGAAGTCAATCTGGAGAAGCTTGGATCCAGTGTCTGAT 501
QY 182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201
Db 502 GGGAACTTCACATTCACCTGCCCATGCCAGATCATCTGTCAGAGTGTGGTTGTGCAAGGCT 561
QY 202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221
Db 562 GTGTCTGTCTGGACATGAGCTCTTCAGAGAGTCCAGTGCACGAGTCTGGCTGAAGAG 621
QY 222 IleLeuCysGlnGlyAsnGlnLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241

```

Db 622 TTCAGTGTGAGGGGAGGAGCCTGAGCTCTGGGTCTGCCACAGAGTGCCTGTCCAGGG 681  
QY 242 HisAspCysSerHisAsnGluAspValThrLeuThrcysThrAspSerSerAspLeuGlu 261  
Db 682 GGCAGGTGCACACAGCTGATCTGTCAGGTGTTGTTGTCACATACACTGAGTCCGG 741  
QY 262 LeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuIleGlnGly 281  
Db 742 CTC--ATGACAACGGCTCTCTCAGTGTGAAGGCGAGTGGAGATTAATCTGGA 798  
QY 282 ArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLys 301  
Db 799 CAATGGAGAGCGCTCTGCTGCCCTCCACTGAGTGTGCCCAATGCAATGTTCTGCT 858  
QY 302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321  
Db 859 CAGCTGGGTGTGGAGTGGCATCTCCACCCCGGAGGACCACTTGGTGGGAAGGT 918  
QY 322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341  
Db 919 GATCAGATCTAACAGCCGATTTCACTGCTCTGGGCTGAGTCCCTCTGTGGATTGT 978  
QY 342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361  
Db 979 CCGTGTGACTGCCCTGGGTGCTGCTGACTGTCATGCCAACAACAGCCTCTGTGATCG 1038  
QY 362 SerAspClyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381  
Db 1039 TCAGGAAC----- 1047  
QY 382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401  
Db 1048 ---CAGATCCAGGTGCTTCCAC----- 1068  
QY 402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421  
Db 1069 -----TGCAACGACTCCGCTG----- 1083  
QY 422 AlalysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441  
Db 1084 TCTCAACTACAGCTCTCGGCGCTCA----- 1110  
QY 442 GluSerAlaLeuTrpAspCysThrThrAspGlyLysAlaLysArgThrCysPheArgArg 461  
Db 1110 ----- 1110  
QY 462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481  
Db 1111 GAGGACAGCGCCCTACTGCTCAGACAGCAGG-----CAGCTCCCGCTGGTGACGGG 1164  
QY 482 HisSerProCysThrGlyArgLeuGluValIlystrGlnGlyGluTrpGlyThrValCys 501  
Db 1165 GGGCGTCCCTGCGCGGAGAGTGGATCCTTGACAGGGCTCTCGGGGACCATCTGT 1224  
QY 502 HisAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGlyLys 521  
Db 1225 GATGACGGCTGGGACCTGGAGATGCCCGGTGTGTGACAGGAGCTGGGCTGTGGAGAA 1284  
QY 522 ProMetHisValPheGlyMetThrThrPheLysGluAlaSerGlyProIleTrpLeuAsp 541  
Db 1285 GCCCTCAATGCCAGGGGTCTGCTCACTTCGGGCGAGGATCAGGGCCCATCTGTTGGAC 1344  
QY 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561  
Db 1345 AACTTGAACGTGCACAGAAAGGAGTCCCATCTGTGGAGGTGCCCTCCCGGGGCTGGGG 1404  
QY 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581  
Db 1405 CAGCACAACGTGCACACACAGAGGAGCGGGGCTCATCTGCTCA-----GAGTTCCTG 1458  
QY 582 GlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValThrPheGln 601  
|||||:|||||

Db 1459 GCCTCAGGATGGTGTGAGTGAGGACACAGCTGCTGGGTGGTGGAGATTTTCTACAAT 1518  
QY 602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCys 621  
Db 1519 GGCACCTGGGGCAGTGTCTGCCGTAAACCCATGGAACACATCACTGTCTCCAGCATCG 1578  
QY 622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640  
Db 1579 AGACAGCTTGGCTGGGGACAGT-----GGAACCCCTCAACTCTCTCTGTTGCT 1626  
QY 641 -----GlyThrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657  
Db 1627 CTTAGAGAAGGTTTATAGCCACACATGGGTGATAGATCCAGTGTCCGAAACACTGACAC 1686  
QY 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677  
Db 1687 TCTCTCTGGCAGTCTCTCTGACCCCTTGAATTAACATCACTGCTCTCCAAAGGAGAA 1746  
QY 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697  
Db 1747 GCCTATATCTGGTGTGCAGACAGC-----AGACAGATCCGCTGGTGGATGGAGTGGT 1800  
QY 698 ArgCysAlaGlyLysValGluValAsnValGlnClyAlaValGlyIleLeuCysAlaAsn 717  
Db 1801 CGTGTCTCTGGAGAGTGGAGATCCCTTGACACGGGCTCTGGGGCACCATTCTGTGATGAC 1860  
QY 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIle 737  
Db 1861 CGTGGGACCTGACAGTGCCTGTGTGTGCAAGCAGCTGGGCTGTGGAGAACCCCTG 1920  
QY 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757  
Db 1921 GAGGCCACTGCTCTCTCTTCTCGGACGGATCAGGGCCCATCTGGCTGGATGAAGTG 1980  
QY 758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777  
Db 1981 AACTGACAGAGAGAGAGTCCCAAGTATGAGGTGCCCTTCTCTGGGATGGCGCACAC 2040  
QY 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797  
Db 2041 AACTGCAATCATCAAGAGATGACAGGATCATCTGCTCAGGATTT-----GTCCGCTG 2094  
QY 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817  
Db 2095 GCTGGAGGAGATGGACCTGCTCAGGCGAGTAGAAGTGCATTCTGGAAGACCTGGACC 2154  
QY 818 SerValCysAspSerAspPheSerHisAlaAlaAsnValLeuCysArgGluLeuAsn 837  
Db 2155 CAGTGTCTGATGAAAACCTTCACTCCCACTCCCACTGCTGTCAGAGCTGGGA 2214  
QY 838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857  
Db 2215 TGTGGCAAGCTGTCTCTGCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGGTC 2274  
QY 858 TrpAlaGlyLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877  
Db 2275 TGGGCTCAAGAGTTCAGGTGTGTGGGGGAGCGCTCAGCTCTGCTGCTGCCCCAGAGTG 2334  
QY 878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897  
Db 2335 CCCTGTCCAGGAGGACATCTCTCCACATGGAGCTCTCAGGTGTCTGTTTCAGTATAC 2394  
QY 898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916  
Db 2395 ACAGAATCCAGTTATGAAAACCGGACCTCTCAATGTGAGGGGAGGTGGAGATGAAG 2454  
QY 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936  
Db 2455 ATCTCTGGACGATGGAGAGCGCTGTGTGCTCCCACTGGAGTCTGGCCAATGCCAATGTT 2514  
QY 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956  
Db 2515 GTCTGTCTCAGTCTGGCTGTGGAGTGCCTCTCCACCCACCCACAGAGACCACTTGGTG 2574







```
Db 4085 TCCCAAACTGGCCATCATGAAGACGCTGCTGCTCATCTCTCAGCTTCCAGTCCAG 4144
Qy 1006 ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGly 1025
Db 4145 CCGACA---CCAGCCAGACACTGGCCCACTGGCCCACTGCATCAACAGCAGGATCTGAA 4201
Qy 1026 SerAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1045
Db 4202 TCCAGTTTG-----GCCCTGAGGCTGGTGAATGGAGGTGCACAGGTGT 4243
Qy 1046 AlaGlyArgValGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1065
Db 4244 CAGGCGGAGTGGAGGTCTATACGAGGCTCCTGGGCGCAGCTGCTGTGATGACTACTGG 4303
Qy 1066 AspLeuSerAspAlaHisValValCysGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1085
Db 4304 GACCAATGATGCAATGTGTTGACGAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4363
Qy 1086 ThrValSerAlaHisPheGlyGluGlySerGlyProLeuLeuLeuLeuLeuLeuLeuLeu 1105
Db 4364 CCAGGAAGTCCCGGTTTGGCCAGGTTTCAGACCCATTGCCCTGGATGATGGCTGTC 4423
Qy 1106 ThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCys 1125
Db 4424 TCAGGACATGATGCTATCTGCTGAGCTGCCGCCCAATGCTGCTGCTGCTGCTGCTGCT 4483
Qy 1126 ArgHisLeuGluAspAlaGlyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1135
Db 4484 GGCCATCATGAAGACGCTGGTGTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4543
Qy 1136 -----SerGluPheThr-----AlaLeu 1141
Db 4544 CCAGACACTTGGCCAACTCTCGTGCATCAACAGCAGGATCTGAATCCACATTGGCCCTG 4603
Qy 1142 ArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsn 1161
Db 4604 AGACTGGTGAAT-----GGAGGTGACAGGTCTCGAGCCGAGTGGAGGCTCTATACCAA 4657
Qy 1162 GlyThrTrpGlySerValGlyArgAsnIleThrAlaIleAlaGlyIleValCys 1181
Db 4658 GCCTCTGGGCGCAGCTGTGTGTGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4717
Qy 1182 ArgGlnLeuLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerTyr 1199
Db 4718 AGGACGCTGGGCTGT-----GGCTGGCCATGTGAGCCGCCAGGAATGCCAGTTT 4768
Qy 1200 ---ThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSer 1218
Db 4769 GCCCAGGCTCAGGACCCCATTTCTCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4828
Qy 1219 IleTrpGlnCysLeuSerAlaProTrp----- 1227
Db 4829 CTGTGGAGCTGCCCCCACTAGTGGCTCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4888
Qy 1228 -----GluArgArgIleSerSerProAlaGluGluThrTrpIleThr 1241
Db 4889 GGTGTCTCTGCTCAGTGTCTCAGTGTCTCAGTGTCTCAGTGTCTCAGTGTCTCAGTGTCT 4948
Qy 1242 -----CysGluAspArgIleArgValArg----- 1249
Db 4949 ACCAACTTACCGGCTTACAGTGTGATCTGAATCCAGTTTGGCTCTCAGGCTGCTGAAT 5008
Qy 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleThrPheHisAlaGlySerTrpGlyThr 1269
Db 5009 GGAGGTGAC---AGGTGTCGAGCGGAGTGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5065
Qy 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnLeuLeuGlyCys 1289
Db 5066 GTGTGTGATGACGCTGGAGACCAATGATCCCAATGATGCTGCTGCTGCTGCTGCTGCTGCT 5125
Qy 1290 GlySerAlaLeuAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1309
```

```
Db 5126 3GCTGGGCGCATGTGGCGCCAGGAATCCCGGTTTGGCCAGGCTCAGGACCCCATTTGTC 5185
Qy 1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329
Db 5186 CTGATGATGTGGCTGCTCAGGGAATGATGCTTACCTTGTGGAGTGGCCGCCCAAGAGGC 5245
Qy 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349
Db 5246 TGGCTCACCCCAACTGTGGCCATCACCAGAGCGTGTGTCTATCTGCTCAGCCACCAAA 5305
Qy 1350 LeuLysSerLeuAsnAlaSerSerGlyHis 1359
Db 5306 ATAAATTTCTACTACGACAGATTGGTGGCAT 5335
RESULT 12
HSA243212 7683 bp mRNA linear PRI 17-SEP-1999
LOCUS Homo sapiens mRNA for DMBT1 protein 8kb transcript variant 2
DEFINITION (DMBT1/8kb.2).
ACCESSION AJ243212
VERSION AJ243212.1 GI:5912463
KEYWORDS DMBT1 gene; DMBT1/8kb.2 protein.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 7683)
AUTHORS Holmskov,U., Mollenhauer,J., Madsen,J., Vitved,L., Gronlund,J.,
Tornoe,I., Kliem,A., Reid,K.B., Poustka,A. and Skjoldt,K.,
Cloning of gp-340, a putative opsonin receptor for lung surfactant
protein D
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (19), 10794-10799 (1999)
MEDLINE 99415938
PUBMED 10485905
REFERENCE 2 (bases 1 to 7683)
AUTHORS Mollenhauer,J.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1999) Mollenhauer J., Molecular Genome Analysis,
Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120
Heidelberg, GERMANY
FEATURES
source
1..7683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q25.3-q26.1"
/tissue_type="trachea"
/dev_stage="adult"
1..7683
/gene="DMBT1"
1..106
/gene="DMBT1"
107..7345
/gene="DMBT1"
/function="putative tumour suppressor, role in immune
system and differentiation"
/notes="Sequence is an alternative splice form of the DMBT1
gene that is expressed in human adult trachea. Isoforms of
DMBT1 are identical to the collectin binding protein
gp-340. Full-length cDNA clone contains 1 bp deletions in
codons 100 and 1751, that were corrected by comparison
with the genomic exons"
/codon_start=1
/product="DMBT1/8kb.2 protein"
/protein_id="CAB56155.1"
/db_xref="GI:5912464"
/translation="MGISTVILEMLLWGVLSGGWIPRTTDYASLPSEVPLDITV
AEGSPFSELTLESTVAEGSPISLESTLETTVAEGSLIPSESTLESTVAEGSDSGAL
RLVNGDCRCQGRVEILYRGSHGAVCDSDNDANDVRCQLGCCWAMSAAPNAPFGQ
SGPITALDDVRCGSHESYLNSCPHNGHSHNGHGDGVCISAAQPSQTLNRPSPVR
ISPPVPTGESSLALRLVNGDCRCGRVEILYRGSHGVTCDYDNDLVNVCRLQ
CGWAMSAAPNAPFGQSGPIVLDVRCGSHESYLNSCPHNGHSHNGHGDGVCIS
APQSRPTSPDPTWPTSTASTAGPESLALRLVNGDCRCQGRVEILYRGSHGVTCDYD
```



QY 454 AlaLysArgThrCysPheArgSerAspAlaGlyValIleCysSer----- 469  
DB 2165 CTCCTCCACAACTGTGCCCATCATGAAGTGTGGTGTCTCTCAGCTGCCCGACGTC 2224  
QY 469 ----- 469  
DB 2225 CGGTGCGAGCCCGCAGACAGCTGTCTGCACCATCATCGTTACCTCATCGACAGTAGGA 2284  
QY 470 AspLysAlaAspLeuAspLeuValGlyAlaHisSerProCysTyrGlyArgLeu 489  
DB 2285 TCTGAATCCAGTTTGACCTTGAGGCTGGTGAATGAGTGCACAGGTGTACGGGCCGAGTA 2344  
QY 490 GluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerTrpArgAsn 509  
DB 2345 GAGTCTCTATACCGAGGCTCTCGGGCACCCTGTGTGATGACAGCTGGGATACCAATGAT 2404  
QY 510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529  
DB 2405 GCCAATGTAGTCTGACGCGAGCTGGCTGTGGTGGCCACGTCGCGCCCGCAGGAATGCC 2464  
QY 530 TyrPheLysGluAlaSerGlyProIleTrpLeuAspValSerCysIleGlyAsnGlu 549  
DB 2465 CGGTTGGCCAGGCTCAGACACCATTTGTTGGATGATGGCTGTCTCAGGACACGAG 2524  
QY 550 SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu 569  
DB 2525 TCCTACCTGTGGAGCTGCCGCCACAATGGTGGCTCTCCACAACTGTGGCCATCATGAA 2584  
QY 570 AspValIleValThrCysSer----- 581  
DB 2585 GATGCTGGTGTCTCATCTGCTCAGTTTCCAGTCCCGCGCAGCACCCAGTCCAGATACTGG 2644  
QY 582 -----GlyLeuArgLeuValGly 587  
DB 2645 CCGACCTCATCATGATCAACAGCAGGACCTGAATCCAGTTGGCCCTGAGGCTGGTGAAT 2704  
QY 588 GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal 607  
DB 2705 GGAGGTGACAGGTGTACAGGCCGAGTGGAGTCTCTATACCGAGGCTCTCGGGCACCGTG 2764  
QY 608 CysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysPro 627  
DB 2765 TGTGATGATAGTGGGACAGTGCAGCCCAATGTGTCTGCCGCGAGCTGGGCTGTGGC 2824  
QY 628 SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeu 647  
DB 2825 TGGCCACGTCAGCCCGCAGGAATGCCGGTTTGGCCAGGTTCCAGACCCATTTGCTCTG 2884  
QY 648 AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp 667  
DB 2885 GATGACGTGGCTGTCTCAGGCTATGATGCTCTACCTGTGGAGTGCCTCCCAACAATGGCTGG 2944  
QY 668 GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla----- 685  
DB 2945 CTCCTCCATAACTGTACGACAGTGAAGAGCGCTGGTGTCTCTCAGCTGCCCACTCC 3004  
QY 685 ----- 685  
DB 3005 TGGTGCAGCCCGCAGCAGACATTCGCGACCATCATCTTGGCTGCATCATGACAGTAGGA 3064  
QY 686 -----SerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysVal 703  
DB 3065 TCTGAATCCAGTTTGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTACGGGCCGAGTG 3124  
QY 704 GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle 723  
DB 3125 GAGGTCTATACCAAGGCTCTCGGGCACCCTGTGTGCGATGACAGCTGGGACACCAATGAT 3184  
QY 724 AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro 743  
DB 3185 GCCAATGTCTGTCAGCGCAACTGGGCTGTGGCTGGGCGCCATGTCTCAGCCCGCAGGAATGCC 3244

QY 744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu 763  
DB 3245 CGGTTGGTCAAGGCTCAGGACCCATGTCTGGATGATGGCGCTGCTCAGGACACGAG 3304  
QY 764 AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet 783  
DB 3305 TCTTACCTGTGGAGCTGCCCCCAACATGGCTGGCTCTCCCAACATGTGGCCATAGTGAA 3364  
QY 784 GluAlaSerLeuIleCysSerAlaHisArg---GlnPro----- 795  
DB 3365 GACCTGGTGTCTCATCTGCTCAGCTTCCAGTCCCGGCCCAACACCTAGTCCAGACACTGG 3424  
QY 796 -----ArgLeuValGly 799  
DB 3425 CCAACCTCATCATGATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 3484  
QY 800 AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal 819  
DB 3485 GGAGTGCACAGGTGTACGGCGCAGTGGAGTCTTATACCAGGCTCTCTGGGGCACCCTGG 3544  
QY 820 CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly 839  
DB 3545 TGTGATGACTACTGGGACACCAATGATGCCAATGTGGCTTCAGGACGCTGGGCTGTGGC 3604  
QY 840 AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAla 859  
DB 3605 TGGGCCATGTGAGCCCGCAGGAATGCCCGGTTTGGCCAGGCTTCAGGACCCCATTTGCTCG 3664  
QY 860 GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis 879  
DB 3665 GATGATGTGGCTGCTCAGGACATGAGTCTATCTGTGGAGCTGCCCCCAACATGGCTGG 3724  
QY 880 ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp 899  
DB 3725 CTCCTCCCAACATGTGGCCATCATGAAGACGCTGGTGTCTCTGCTCAGCTTCCAGTCC 3784  
QY 899 ----- 899  
DB 3785 CAGCCGACACCCCGCAGACACTTGGCCAACTTCACATGCATCAACAGCAGGATCTCAA 3844  
QY 900 -----ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle 915  
DB 3845 TCCAGTTTGGCCCTCAGGCTGGTGAATGGAGGTGACAGGTGTACGGCGCAGTGGAGTGC 3904  
QY 916 AsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArg 935  
DB 3905 CTATACCGAGGCTCTCGGGCACCGGTGTGTGATGACTACTGGACACCAATGATGCCAAT 3964  
QY 936 ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyLysTyrIle 955  
DB 3965 GTGTTTGCAGGACGCTGGGCTGGTGGCCACGTCAGCCAGCCAGGAATGCCCGGTTT 4024  
QY 956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975  
DB 4025 GGCCAGGTTTCAGGACCCATTTGCTCTGGATGATGTGGCTGTCTCAGGACATGAGTCTAT 4084  
QY 976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrVal 995  
DB 4085 CTGTGGAGTGCCTCCCAACATGGGCTCTCCCAACATGTGGCCATCATGAAGACGCT 4144  
QY 996 SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer 1015  
DB 4145 GGTGTCTATCTCTCAGCTTCCAGTCCAGTCCCGCAGCA---CCAGCCACAGACACTGGCCA 4201  
QY 1016 AspProTrpLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg 1035  
DB 4202 ACCTCATGATCAACAGCAGGATCTGAATCCAGTTTG-----GCC 4243  
QY 1036 LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly 1055  
DB 4244 CTGAGGCTGGTGAATGAGGTGACAGGTGTCTCAGGCCGCGAGTGGAGGTCTTATACGAGGC 4303  
QY 1056 PheTrpGlyThrIleCysAspGlyTrpAspLeuSerAspAlaHisValValCysGln 1075

Db 4304 TCCTGGGGACCGTGTGATGACTACTGGACACCAATGATCCCAATGCTGTTTCGAGG 4363  
QY LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer 1095  
Db 4364 CAGCTGGCTGTGGCTGGCCAGCTGAGCCAGGAAATGCCGGTTTGGCCAGGTTTCA 4423  
QY GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCys 1115  
Db 4424 GGACCAATTGCTGATGATGCGCTGCTCAGGACATGAGTCCCTATCTGTGGAGCTGC 4483  
QY ProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCys 1135  
Db 4484 CCCACAATGGCTGGCTCTCCCAACAATGTGGCCATCATGAAGACGCTGCTCATCTGC 4543  
QY SerGluPhe 1138  
Db 4544 TCAGCTTTCAGTCCAGCCGACACCCAGCCAGACACTTGCCCAACCTCTCGTCATCA 4603  
QY 1139 -----ThrAlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151  
Db 4604 ACAGCAGGATCTGAATCCACTTTGGCCCTGAGACTGGTGAAT-----GGAGGTGACAGG 4657  
QY CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsn 1171  
Db 4658 TGTCGAGCGGAGTGAGGTCTCTATACCAAGGCTCTGGGGACCGTGTGTGATGACTAC 4717  
QY 1172 IleThrAlaLeuAlaGlyLeuValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191  
Db 4718 TGGGACACCAATGATGCCAAGTGGTCTGCAGCAGCTGGGCTGT-----GGCTGG 4768  
QY 1192 ValSerLeuAlaProLeuSerLys-----ThrGlySerGlyPheMetTrpValAsp 1208  
Db 4769 GCATGTCAGCCGACCAAGAAATGCCAGTTTGGCCAGGCTCAGGACCCATCTCTCGGAT 4828  
QY 1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227  
Db 4829 GATGTCGCTGCTCAGGACACGAGCTTACCTGTGGAGTGGCCCAATGGCTGGCTC 4888  
QY 1228 -----GluArgArgIle 1231  
Db 4889 TCCACAACATGTGGCCATCATGAAGATGCTGTGTCTATCTGCTCAGTCTCAGTCCACAG 4948  
QY 1232 SerSerProAlaGluGluThrTrpIleThr-----GlyGlyAspThrGluCysSerGlyArgVal 1259  
Db 5009 GAATCCAGTTTGGCTGAGGCTGGTGAATGAGGTGAC---AGGTGTCGAGCGCCAGTG 5065  
QY 1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGlu 1279  
Db 5066 GAGGTCTCTATCGAGGCTCTCGGGGAACCGTGTGTGATGACAGCTGGGACACCAATGAT 5125  
QY 1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAla 1299  
Db 5126 GCATGTGCTGTCAGGACGCTGGCTGTGGTGGCCATCTGCGGCTCAGGCGCCAGAAATGCC 5185  
QY 1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319  
Db 5186 CGGTTTGGCCAGGCTCAGGACCCATTTGCTGGGTGATGTGCTGCTCAGGGAATGAG 5245  
QY 1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339  
Db 5246 TCTTACCTGTGGAGCTGGCCGACCAAAAGGCTGGCTCACCCCACTGCTGGCCATCACGAA 5305  
QY 1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359  
Db 5306 GAGCTGGTGTCTATGCTCAGCCACCAATAAATTTACTACGACAGATGGTGGCAT 5365  
RESULT 13  
HSA297935

LOCUS HSA297935 7667 bp mRNA linear PRI 12-JAN-2002  
DEFINITION Homo sapiens mRNA for DMBT1 8 kb transcript variant 2 (DMBT1 gene).  
ACCESSION AJ297935  
VERSION AJ297935.1 GI:14715230  
KEYWORDS alternative splicing; deleted in malignant brain tumors 1; DMBT1 gene; DMBT1/8kb.2 protein.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
Mollenhauer, J., Herbertz, S., Helmke, B., Kollender, G., Krebs, I., Madsen, J., Holmskov, J., Sorger, K., Schmitt, L., Wiemann, S., Otto, H. F., Grone, H. J. and Poustka, A.  
Deleted in Malignant Brain Tumors 1 is a versatile mucin-like molecule likely to play a differential role in digestive tract cancer  
Cancer Res. 61 (24), 8880-8886 (2001)  
JOURNAL MEDLINE 21621923  
PUBMED 11751412  
REFERENCE 2 (bases 1 to 7667)  
AUTHORS Mollenhauer, J.  
TITLE Direct Submission  
JOURNAL Submitted (09-NOV-2000) Mollenhauer J., Molecular Genome Analysis, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120 Heidelberg, GERMANY  
FEATURES  
Location/Qualifiers  
1..7667  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10q25.3-q26.1"  
/tissue\_type="small intestine"  
/dev\_stage="adult"  
88..7329  
/gene="DMBT1"  
88..7329  
/gene="DMBT1"  
/function="putative role in mucosal protection, immune response, wound healing"  
/note="DMBT1 8kb transcript variant 2 deleted in malignant brain tumors 1 alternative splice variant, contains all exons except for exon 1"  
/codon\_start=1  
/evidence=experimental  
/product="DMBT1/8kb.2 protein"  
/protein\_id="CAC44122.1"  
/db\_xref="GI:14715231"  
/translation="MGISTVILEVCLLWQVLTGGWIPRTTDTIASLIPSEVLPDPTV  
AEGFPFSESTLESTVAEGSPISLESTLESTVAEGSLIPSESTLESTVAEGSDSLAL  
RLVNGDRCGRVEILYRGSGTVCDDSDTNDANVYRQLCGCWAMAPGNARFGQSGP  
SGPTALDDVRCSCGHELYLWSCPNGWLSHNGHGEDAGVICSAAPQSTLTPESWVPR  
ISPPVPTGESSIALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANVYRQLCG  
CWAMAPGNARFGQSGPILVDDVRCSCGHELYLWSCPNGWLSHNGHGEDAGVIC  
APQSRPTSPDTWPTSHASTAGPESLALRLVNGDRCGRVEILYRGSGTVCDDSD  
DTNDANVYRQLCGCWAMAPGNARFGQSGPILVDDVRCSCGHELYLWSCPNGW  
NCHSDAGVICSAHRSPTSPDTLITLPASTGSESSIALRLVNGDRCGRVEILY  
VLVYRGSGTVCDDSDTNDANVYRQLCGCWAMAPGNARFGQSGPILVDDVRCSG  
TVCCDDSDTNDANVYRQLCGCWAMAPGNARFGQSGPILVDDVRCSCGHELYLW  
NGLWSHNGHGEDAGVICSAHRSPTSPDTLITLPASTGSESSIALRLVNGDRC  
RCGRVEILYRGSGTVCDDSDTNDANVYRQLCGCWAMAPGNARFGQSGPILV  
DVRCSCGHELYLWSCPNGWLSHNGHGEDAGVICSAHRSPTSPDTWPTSHAST  
ESSIALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANVYRQLCGCWAMAPGN  
ARFGQSGPILVDDVRCSCGHELYLWSCPNGWLSHNGHGEDAGVICSAHRSPT  
DTLPTLTPASTGSESSIALRLVNGDRCGRVEILYRGSGTVCDDSDTNDANVY  
CRLCGCWAMAPGNARFGQSGPILVDDVRCSCGHELYLWSCPNGWLSHNGHSE  
GVTCASASQSRPTSPDTWPTSHASTAGSESSIALRLVNGDRCGRVEILYRGSG  
CDYWDTNDANVYRQLCGCWAMAPGNARFGQSGPILVDDVRCSCGHELYLWSCP  
NGLWSHNGHGEDAGVICSAHRSPTSPDTWPTSHASTAGSESSIALRLVNGDRC  
RVEILYRGSGTVCDDSDTNDANVYRQLCGCWAMAPGNARFGQSGPILVDDVRC  
SGHELYLWSCPNGWLSHNGHGEDAGVICSAHRSPTSPDTWPTSHASTAGSESS

ALRLVNGDRGCRGRVLYRGSGVTGDDYWDNDANVVCRLGCGWATSAPONAREG  
 QSGSPVLDDVDSRGHESYLWSCPHNGHSHNGHEDAGVICSASQSPDTPSPDTP  
 TSTRAGTSESTLALRVNGDRGRVLYRGSGVTGDDYWDNDANVVCRLGCG  
 GWASAPNAQFGSGSPVLDDVDSRGHESYLWSCPHNGHSHNGHEDAGVICS  
 AQOSQTPRDTWLTNLPAITVGSSESLALRVNGDRGRVLYRGSGVTGDDV  
 WDNDANVVCRLGCGWASAPNAREFGSGSPVLDDVDSRGHESYLWSCPHNGWLT  
 HNGHEDAGVICSATQINSTTDDWHPTTTTARESSNCGGFLFYASGTFSSYP  
 YPNNAKCVWEIEVNSGYRNLGFSNLKLEAHNCSFDYVEIFDGLSNLSLLIKTCN  
 IYHGFTWGTVCDDSWTIOAEVVCRLGCGRAVSNALNAYFGSGSPITLDDVCSGT  
 ESTLWOCNRGHSNCHREDAGVICSNHLSTPAFFLNITRPNDYSCGGFLSQPS  
 GFSPFPGENTPNNAKCVMDIEVQNNYRVTIFRDVLEGGCNDYIEVFDGPIESS  
 PLIARCDGARSSTSSNMFIRFISDSHITRRGEAYEYSSPNSDNLCLPNHM  
 QASVSYSYQSLGFSADLVISTWNGYIECRPOITNLVIFIPYSGCGTFKQADNT  
 IDYSNLTAAVSGGILKRTDLRIHVSCRMLONTWDTMYIANDTTHVANNTIQVEV  
 QYGNEDVNTSFTSSSFLYPTVRPYVDLNOGLVVOAELHSDAVLTLEVDTCVAP  
 YSNDFSLIYDLIRSCVRDDTYGPYSSPLRIARFRPRAFLHFLNPPSYLHCKMNV  
 CRAIDPSSKICRGCVLRKSRDVGSGYQEKVDVVLGPILQLOTAPRREPR"

sig\_peptide

mat\_peptide

BASE COUNT 1605 a 2137 c 2166 g 1759 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 6.39e-185 Length: 7667  
 Score: 2982.00 Matches: 618  
 Percent Similarity: 51.56% Conservative: 176  
 Best Local Similarity: 40.13% Mismatches: 506  
 Query Match: 36.64% Indels: 240  
 DB: 9 Gaps: 23

US-09-759-130B-381 (1-1453) x HSA297935 (1-7667)

QY 44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63  
 DB 775 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGCGACAGGTGTCGAGCGCGAGTGGAGTGC 834  
 QY 64 LysPheGlnGlyGlnTrpGlyThrValCysAspGlyTrpAsnThrAlaSerThr 83  
 DB 835 CTATACCGAGGCTCCTGGGCGACCGTGTGTGATGACTTGGGACCAATGATGCCAAT 894  
 QY 84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet-----PheArgPhe 100  
 DB 895 GTGCTGTCAGGAGCTGGGCTGGCTGGCCATGTCAGCCCCAGGAAATGCCAGTTT 954  
 QY 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspValSerCysTyrGlyAsn 120  
 DB 955 GGCCAGGCG-----TCAGGACCACTTGTCTCGATGATGCTGCTCAGGACAC 1005  
 QY 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140  
 DB 1006 GAGTCTTACCTGTGAGCTGCCCAATGAGTGGCTGCCTACCCACAATGTGGCCATAGT 1065  
 QY 141 GluAspValGlyValAsnCysTyrGly----- 149  
 DB 1066 GAAGACGCTGGTGTCTGCTCAGCTCCCGAGTCCCGCGCGACACCCAGCCAGATACT 1125  
 QY 150 -----GluAlaAsnLeuGlyLeuArgLeuVal 158  
 DB 1126 TGGCCGACCTCACATGATCAACAGCAGGACCTGAATCCAGTTTGGCCCTCAGGCTGGT 1185  
 QY 159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178  
 DB 1186 AATGAGGTGACAGGTGTCAGGCGCGAGTGGAGTCTTATACCGAGGCTCCTGGGGCACC 1245  
 QY 179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValCysArgGlnLeuGlyCys 198  
 DB 1246 GTGTGTATATAGTCTGGGACACCATGACGCCAATGTGTCTGCCGCGAGCTGGGCTGT 1305  
 QY 199 ProSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218

DB 1306 GGCTGGCCACGTCAGCCCGCAGGAAATGCGCGTTCGCCAGGTTTCAGGACCCATTGTC 1365  
 QY 219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238  
 DB 1366 CTGATGACGTGGCTGCTCAGGCTATGAGTCTTACCTGTGGAGCTGCCCCACAAATGGC 1425  
 QY 239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyr----- 255  
 DB 1426 TGGCTCTCCCACTACTGTCAGCACAGCTGAAGACCGCTGGTGTCTCTGCTCAGCTGCCAC 1485  
 QY 255 ----- 255  
 DB 1486 TCCAGGTCGACGCCACGTCCTCCAGACACCGTTCGCCAGCATCACCTTACCTGTCATCGACAGTA 1545  
 QY 256 ---AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArg 274  
 DB 1546 GGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGAGGTGACAGGTGTACAGGGCGGA 1605  
 QY 275 ValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAla 294  
 DB 1606 GTGAGGTCTATATACCGAGGCTCCTGGGCGACCGTGTGTATGACAGCTGGGACACCAAT 1665  
 QY 295 AlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeu 314  
 DB 1666 GATGCCAATGTGTCTCAGGCGAGCTGGGCTGTGGCTGGCCATGTGGCCCCAGGAAT 1725  
 QY 315 ProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsn 334  
 DB 1726 GCCCGTTTGGTCAGGCTCAGGACCATGTCTCTGATGATGCTGCGCTCAGGGAAT 1785  
 QY 335 GluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGln 354  
 DB 1786 GAGTCTTGTGTGAGCTGCCCGCACCAATGGCTGGCTCTCCCAATACTGTGGCCATAGT 1845  
 QY 355 AsnAspValSerValIleCysSer---AspGlyAlaAspLeuGluLeuArgLeuAlaAsp 373  
 DB 1846 GAAGACGCTGGTGTCTCAGGACCTGAATCCAGTTTGGCCCTCAGGCTGGTGAAT 1905  
 QY 374 GlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrpTrpIle 393  
 DB 1906 GGAGTGCAGGTGTCAGGCGCGAGTGGAGTCTTATACCGAGGCTTGTGGGGCACCGTG 1965  
 QY 394 CysAspGlnAsnTrpLysAsnGluAlaLeuValValCysLysGlnLeuGlyCysPro 413  
 DB 1966 TGTGATGACAGCTGGGACACCAATGATGCCAATGTGTCTCAGGCGAGCTGGGCTGGC 2025  
 QY 414 PheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIle 433  
 DB 2026 TGGGCCACGTCAGCCCCAGGAAATGCCCGTGTGGTGCAGGCTCAGGACCCATTGCTCTG 2085  
 QY 434 AsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLys 453  
 DB 2086 GACATGTGGCTGCTCAGGACATGAGTCTTACCTGTGGAGCTGCCCCACAAATGGCTGG 2145  
 QY 454 AlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSer----- 469  
 DB 2146 CTCTCCCAACAATGTGCGCATCATGAAGATGCTGGTGTCTCATCTGCTCAGCTGCCAGTCC 2205  
 QY 469 ----- 469  
 DB 2206 CGGTGACGCCCGCGGCGAGACACGTTGTGACCATCATCGTTTACCTCCATCGACAGTAGGA 2265  
 QY 470 AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeu 489  
 DB 2266 TCTGAATCCAGTTTGACCTTGAGGCTGGTGAATGAAGTGCAGGCTGTACAGGCGCGAGTA 2325  
 QY 490 GluValLysTrpGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsn 509  
 DB 2326 GAGTCTCTATACCGAGGCTCCTGGGCGACCGTGTGTGATGACAGCTGGGATACCAATGAT 2385  
 QY 510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529

Db	2386	GCCAATGTGCTGTCAGGAGCGCTGGCGCTGTGCTGGCGCCACGTCGGCGCCACGAAATGCC	2444
Qy	530	TyrPhelysGluAlaSerGlyProIleTrpLeuAspValSerCysIleGlyAsnGlu	549
Db	2446	CGGTTTGGCCAGGCGTCAGGACCCATTGCTTGGATGATGTGCGCTGCTCAGGACACAG	2505
Qy	550	SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu	569
Db	2506	TCCTTACCTGTGGAGTGCCTCCACAAATGCTGGCTCTCCCACTGTGGCCATCATGAA	2565
Qy	570	AspValIleValThrCysSer-----GlyAspAlaThrTrp	581
Db	2566	GATGCTGGTGTCACTGCTCAGTTTCCCACTCCCGCGCGACACCCAGTCCAGATACTGG	2625
Qy	582	-----GlyLeuArgLeuValGly	587
Db	2626	CCGACCTCACATGCATCACAGGAGTCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT	2685
Qy	588	GlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrVal	607
Db	2686	GCAGGTCACAGGTGCTCAGGCGCGAGTGGAGGTCTCTATACCGAGGCTCTCTGGGCGCACCG	2745
Qy	608	CysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysPro	627
Db	2746	TGTGATGATAGCTGGGACACAGTAGCAGCGAAATGTGTCTCCGCGCAGCTGGGCTGGC	2805
Qy	628	SerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeu	647
Db	2806	TGGCCACGTCAGCCCGAGGAATGCCCGTTTGGCCAGGTTTCAGGACCCATTGTCCTG	2865
Qy	648	AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp	667
Db	2866	GATGACGTGCGCTGCTCAGGCTATGAGTCTACCTGTGGAGCTGCCCGCACAAATGGCTGG	2925
Qy	668	GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla	685
Db	2926	CTCTCCCATAACTGTCAGCACAGTGAAGACGCTGGTGTCTATGCTCAGCTGCCCACTCC	2985
Qy	685	-----	685
Db	2986	TGCTGAGGCCGACATCGACACATGCGCGACCATCAGTTGCTGCTGCATCGACAGTAGGA	3045
Qy	686	-----SerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLysVal	703
Db	3046	TCTGAATCCAGTTTGGCCCTGAGGCTGGTGATGAGGAGTGCAGAGTGTACAGGCGCGAGTG	3105
Qy	704	GluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIle	723
Db	3106	GAGTCTCATACCAAGGCTCTGGGCGACCGTGTGCGATGACAGTGTGGACACCAATGAT	3165
Qy	724	AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro	743
Db	3166	GCCAAATGCTGTCGAGGCAACTGGGCTGTGGCTGGCGCATGTGAGCCCGCAGGAAATGCC	3225
Qy	744	HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGlu	763
Db	3226	CGGTTTGGTCAGGCGTCAGGACCCATTGCTCGGATGATGTCGCTGCTCAGGACACAG	3285
Qy	764	AlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMet	783
Db	3286	TCCTTACCTGTGGAGTGCCTCCCAATGGCTGCTCTCCCACTGTGGCCATAGTGAA	3345
Qy	784	GluAlaSerLeuIleCysSerAlaHisArg---GlnPro-----	795
Db	3346	GAGCTGGTGTCACTGCTCAGTTCCCACTCCCGCGCCACACCTAGTCCACAGCACTGG	3405
Qy	796	-----ArgLeuValGly	799
Db	3406	CAAACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGCTGAAT	3465
Qy	800	AlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerVal	819
Db	3466	GGAGGTGCACAGGTGTGAGGCGCGAGTGGAGTCTCTATACCGAGGCTCTCTGGGCGCACCG	3525

QY	820	CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly	839
DB	3526	TGTGAGTACTGGGACACCAATGATGCCAATGTGTGTTGCAGGACGCTGGGTGGC	3585
QY	840	AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrrPala	859
DB	3586	TGGGCCATGTCAGCCCCAGGAATGCCGGGTTTGGCCAGGGTTTCAGGACCCATGTGCTCTG	3645
QY	860	GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis	879
DB	3646	GATGATGTGGCTGCTCAGGACATGAGTCTATCTGTGGAGCTGCCCCCAACATGGCTGG	3705
QY	880	ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp	899
DB	3706	CTCTCCCAACTGTGGCCATCATGAAGACGCTGGTGTCTATCTGCTCAGCTGCCAGTCC	3765
QY	899	-----	899
DB	3766	CAGCCGACACCCAGCCAGACACTTGGCCAACTCATCATCATCAACAGGAGCTGAA	3825
QY	900	-----ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle	915
DB	3826	TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTCTCAGGCCCGAGTGGAGTCT	3885
QY	916	AsnValLeuGlyHisThrPglySerLeuCysAspThrHisThrAspSerProGluAspAlaArg	935
DB	3886	CTATACCGAGGCTCTGGGACCCGGTGTGTGATGACTACTTGGGACACCAATGATGCCAAT	3945
QY	936	ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIle	955
DB	3946	GTGGTTTGCAGGCAGCTGGGCTGTAGCTGGGCCACGCTGACCCCGAGAAATGCCGGTTT	4005
QY	956	GlyGluArgSerValArgValTrrPglyHisArgPheHisCysLeuGlyAsnGluSerLeu	975
DB	4006	GCCAGGGTTCAGGACCCATTTGCTCTGGATGATGTGGCTGCTCAGACATGAGTCCAT	4065
QY	976	LeuAspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrVal	995
DB	4066	CTGTGGAGCTGCCCCACCAATGCCTGGTCTCCCACTGTGGCCATCATGAAGCGCT	4125
QY	996	SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer	1015
DB	4126	GGTGTATCTGTCTCAGCTCCCACTGCCAGTCCAGCAGACA---CCAGCCACACACTTGGCCA	4182
QY	1016	AspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg	1035
DB	4183	ACCTCATATGCATCACAGCAGGATCTGAATCCAGTTG-----GCC	4224
QY	1036	LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly	1055
DB	4225	CTGAGGCTGGTGAATGGAGGTGACAGGTGTCTCAGGCCGCGAGTGGAGGTCTCTATACCGAGGC	4284
QY	1056	PheTrrPglyThrIleCysAspAspGlyTrrAspLeuSerAspAlaHisValValCysGln	1075
DB	4285	TCTTGGGGACCGGTGTGTATGACTACTGGGACCACTGATGCCAATGCTTGGTTTGGCAGG	4344
QY	1076	LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer	1095
DB	4345	CAGCTGGGCTGTGGCTGGGCCACGTCAGCCCAAGGAAATGCCCGGTTTGGCCAGGTTCA	4404
QY	1096	GlyProIleTrrPLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrrPglyCys	1115
DB	4405	GGACCATTTGCTGTGATGTGGCTCTGCTCAGGACATGAGTCCATCTCTGTGGAGCTGC	4464
QY	1116	ProSerArgGlyTrrPglyGlnHisAspCysArgHisLysLysAspAlaGlyValIleCys	1135
DB	4465	CCCCAATGGCTGGCTCTCCACACACTGTGGCCATCATGAAGCGCTGTGTCTCATCTGC	4524
QY	1135	-----	1135
DB	4525	TCAGTTTCCAGTCCCGAGCGGCACACCCAGACACTTGGCCAACTTCTCGTGCATCA	4584

Qy 1136 -----SerGluPheThr-----AlaLeuArgLeuTyrSerGluThrGluThrGluSer 1151  
 Db 4585 ACAGCAGGATGAATCCACTTGGCCCTGAGACTGTGTAAT-----GGAGGTGACAGG 4638  
 Qy 1152 CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgAsn 1171  
 Db 4639 TGTGAGCGCGAGTGGAGGTCCTATACCAAGGCTCCGTGGGACACCGTGTGTGATGACTAC 4698  
 Qy 1172 IleThrThrAlaLeuAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191  
 Db 4699 TGGGACACCAATGATGCCACAGTGTCTGAGCAGCAGCTGGGCTGT-----GGCTGG 4749  
 Qy 1192 ValSerLeuAlaProLeuSerLys-----ThrGlySerGlyPheMetTrpValasp 1208  
 Db 4750 GCCATGTGACGCCCCCAGGAATGCCAGTGTGGCCAGGCTCAGACCCATGTCTGGAT 4809  
 Qy 1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227  
 Db 4810 GATGTGCGCTGCTCAGGACAGACTTACCTGTGGAGCTGCCGCCACAAATGGCTGGCTC 4869  
 Qy 1228 -----GluArgArgile 1231  
 Db 4870 TCCCAACATGTGCCCATCATGAAGTGGTGTCTCATCTCCTCAGCTGCTCAGTCCAG 4929  
 Qy 1232 SerSerProAlaGluGluThrTrpIleThr-----Cys 1242  
 Db 4930 TCAACGCCAGCCAGATGATCTGGCTGACCAACCACTTACCGGCTTACAGTAGGATCT 4989  
 Qy 1243 GluAspArgileArgValArg-----GlyGlyAspThrGluCysSerGlyArgVal 1259  
 Db 4990 GAATCCAGTTGGCTCTGAGGCTGGTGAATGGAGGTGAC---AGGTGTCAGAGCCGAGTG 5046  
 Qy 1260 GluIleTrpHisAlaGlySerTrpTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279  
 Db 5047 GAGTCTCTGATCGAGGCTCTCGGGAACCGTGTGTGATGACACTGGGACACCAATGAT 5106  
 Qy 1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAla 1299  
 Db 5107 GCCAATGTGTCTGACGAGCTGGGCTGGTGGCCATGTGCGGCCACAGGAATGCC 5166  
 Qy 1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319  
 Db 5167 CGGTTTGGCCAGGCTCAGACACCATGTCTGATGATGATGCTGCTCAGGGAATGAG 5226  
 Qy 1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339  
 Db 5227 TCCTACCTGTGGAGCTGCCCCCAACAAAGGCTGGCTCACCACCACTGTGCCCATCAGAA 5286  
 Qy 1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359  
 Db 5287 GAGCTGTGTCTGATCTGCTCAGCCACCAATAATCTACTACGACAGATTTGGTGGCAT 5346

RESULT 14  
 AF159456 7686 bp mRNA linear PRI 16-SEP-1999  
 LOCUS Homo sapiens gp-340 variant protein (DMBT1) mRNA, complete cds.  
 DEFINITION AF159456  
 ACCESSION AF159456  
 VERSION AF159456.1 GI:5733597  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 7686)  
 Holmskov, U., Mollenhauer, J., Madsen, J., Vitved, L., Gronlund, J.,  
 Tornoe, I., Kliem, A., Reid, K.B., Poustka, A. and Skjodt, K.  
 Cloning of gp-340, a putative opsonin receptor for lung surfactant  
 protein D  
 Proc. Natl. Acad. Sci. U.S.A. 96 (1999):10794-10799 (1999)  
 JOURNAL MEDLINE 99415938  
 PUBMED 10485905  
 REFERENCE 2 (bases 1 to 7686)  
 Holmskov, U., Mollenhauer, J., Madsen, J., Vitved, L., Gronlund, J.,

Tornoe, I., Kliem, A., Reid, K.B.M., Poustka, A. and Skjodt, K.  
 Direct Submission  
 Submitted (16-JUN-1999) Department of Immunology & Microbiology,  
 Institute of Medical Biology, University of Southern Denmark,  
 Odense University, Winsloe Parken 21.1, Odense 5000, Denmark  
 Location/Qualifiers  
 1. 7686  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /map="10q25.3-26.1"  
 1. 7686  
 /gene="DMBT1"  
 /note="deleted in malignant brain tumor 1"  
 107. 7348  
 /gene="DMBT1"  
 /note="scavenger receptor cysteine-rich protein SRCR;  
 putative receptor for gp-3"  
 /codon\_start=1  
 /product="gp-340 variant protein"  
 /protein\_id="A049696.1"  
 /db\_xref="GI:5733598"  
 /translation="MGISTVILEMCLLWGVLTGGWIPTTDTASLIPSEVPLDQTV  
 AEGSPFSESTLESTAEAGSPIISLESTLESTVAEGSLIPSESTLESTVAEGSGGLAL  
 RLNGDRCQGRVEILYRGSGTVCDDSDMTNDANVYVQLGCGWASAPGNAMFGQ  
 SGPIALDDVRCSESSYLWSCPHNGWLSHNGHGEDAGVICSAAQPOSTLRPESWVR  
 ISPVTEGSESSYLALRLVNGDRCGRVEILYRGSGTVCDDSDMTNDANVYVQLG  
 CGWASAPGNAQFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHGEDAGVICS  
 APGRTPSPDTPWPTSHASTAGPESLALRLVNGDRCGRVEILYRGSGTVCDDSD  
 TSDANVYVQLGCGWATAPGNARFGQSGPIVLDVRCSESSYLWSCPHNGWLSH  
 NCQHSAGVICSAAHSMWPTPTITLPASTVSESSYLALRLVNGDRCGRVE  
 LYLRGSGTVCDDSDMTNDANVYVQLGCGWAMLPAGNARFGQSGPIVLDVRCSE  
 ESYLWSCPHNGWLSHNGHGEDAGVICSPESSLALRLVNGDRCGRVEILYRGSG  
 TVCDDSDMTNDANVYVQLGCGWASAPGNARFGQSGPIVLDVRCSESSYLWSCP  
 NGWLSHNGHGEDAGVICSAAQSRTPPTDITLTPSTVSESSYLRLVNGSD  
 RCGRVEILYRGSGTVCDDSDMTNDANVYVQLGCGWASAPGNARFGQSGPIVLD  
 DVRCSESSYLWSCPHNGWLSHNGHGEDAGVICSQSRTPPTDTPWPTSHASTAGS  
 ESSLALRLVNGDRCGRVEILYRGSGTVCDDSDMTNDANVYVQLGCGWATAPGN  
 ARFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHGEDAGVICSAAHSMWPTSP  
 DTLPITLPASTVSESSYLALRLVNGDRCGRVEILYRGSGTVCDDSDMTNDANV  
 CROPGCGWASAPGNARFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHSEDA  
 GVICSASQSRTPPTWPTSHASTAGSESSLALRLVNGDRCGRVEILYRGSGTVC  
 CDDSDMTNDANVYVQLGCGWASAPGNARFGQSGPIVLDVRCSESSYLWSCPHN  
 GWSHNGHGEDAGVICSASQSRTPPTWPTSHASTAGSESSLALRLVNGDRCGR  
 RVBLYRGSGTVCDDSDMTNDANVYVQLGCGWATAPGNARFGQSGPIVLDVRC  
 SGHESYLWSCPHNGWLSHNGHGEDAGVICSASQSRTPPTWPTSHASTAGSESS  
 ALRLVNGDRCGRVEILYRGSGTVCDDSDMTNDANVYVQLGCGWATAPGNARFG  
 QSGPIVLDVRCSESSYLWSCPHNGWLSHNGHGEDAGVICSASQSRTPPTWPT  
 TSRASTAGSESSLALRLVNGDRCGRVEILYRGSGTVCDDSDMTNDANVYVQLG  
 CGWASAPGNAQFGQSGPIVLDVRCSESSYLWSCPHNGWLSHNGHGEDAGVICS  
 AQSQSRTPPTWPTTLNLPALTVSESSLALRLVNGDRCGRVEILYRGSGTVCDD  
 SDMTNDANVYVQLGCGWASAPGNARFGQSGPIVLDVRCSESSYLWSCPHNGW  
 LSHNGHGEDAGVICSATQINSTTDMWHTTTTARPSNCGGFLFYASGTFSSPYA  
 HNCGHEDAGVICSATQINSTTDMWHTTTTARPSNCGGFLFYASGTFSSPYA  
 YYPNNAKCVIEIVSGYRINLGFSLNLEAHNCSEDFYVEIFDGLSNLSLLGKICN  
 DTROITSSNMTIHFSDIFONTGLAWNSFPSSDALRLVNLNNSYLKACGRVE  
 IYHGWTGTVCDDSDMTIQEAEVYVQLGCGWASAPGNARFGQSGPIVLDVRCSE  
 ESTLWCCNRGFWFHNCHNEDAGVICSNHLSTPAPFLNITRPNTDYSCGGLSOPS  
 GDFSPPYPCNYPNNAKCVMDIEQNNYRVTVIFRDVLGCGGNYDIEVDFDGPYRS  
 PLIARVCDGARGSTSSNEMSTFISDHSITRGRFAEYVSSPNDSNTLLCLPNHM  
 QASVRSYLSLGSFASDLVSTWNGYECRPQITPNLVIPTIPIYSCGTFKQNDT  
 IDYSNLFATAASGGIIRKRDRLRIHSCRMQLTNDWTDYTIANDTHAVANTIQEVE  
 QYGFNDYNIISFTSSSFLPVTSPRYVLDNOLYQAEILHSDAVLFLRVDTCVASP  
 YSNDFTSLYDLIRSGCRDDTYGPPSSPLIRAFRFAHFLNRPSPVYLRCKMNV  
 CRAYDPSRCYRCGLRSKRKDVGSYQEKVDVLGPIQLQIQTTPRREPR"

BASE COUNT 1614 a 2141 c 2168 g 1763 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.66e-185 Length: 7686  
 Score: 2980.00 Matches: 618  
 Percent Similarity: 51.62% Conservative: 177  
 Best Local Similarity: 40.13% Mismatches: 505  
 Query Match: 36.62% Indels: 240

DB: 9 Gaps: 23  
US-09-759-130B-381 (1-1453) x AF159456 (1-7686)  
Qy 44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63  
Qy 794 TCACATTGGCCCTGAGCTGGTGAATGAGCGGACAGGTGTCCGAGCGGAGTGAGGTC 853  
Qy 64 LysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrAlaSerThr 83  
Db 854 CTATACCGAGGCTCTGGGGACCGCTGTGTGATGACTACTGGGACACCAATGATGCCAAT 913  
Qy 84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet-----PheArgPhe 100  
Db 914 GTGGCTGCAGGACGCTGGCTGGTGGCCATGTCTGGATGATGTGGCTGCTCAGGACAC 973  
Qy 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTrpGlyAsn 120  
Db 974 GCGCCAGGCG-----TCAGGACCCCATGTCTGGATGATGTGGCTGCTCAGGACAC 1024  
Qy 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTrpHisGly 140  
Db 1025 GAGTCTCTACCTGTGAGCTGCCCCACAAATGGCTGGCTCACCACCACTGTGGCCATAGT 1084  
Qy 141 GluAspValGlyValAsnCysTrpGly----- 149  
Db 1085 GAAGACGCTGGTGTCTATCTGCTCAGCTCCCGAGTCCCGCGCGACACCCAGCCAGATACT 1144  
Qy 150 -----GluAlaAsnLeuGlyLeuArgLeuVal 158  
Db 1145 TGGCCGCACTCATATGCATCAACACGAGGACCTGAATCCAGTTTGGCCCTGAGGCTGGT 1204  
Qy 159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178  
Db 1205 AATGAGGTGACAGGTGTGAGCGGCGAGTGGAGTCTTATACCGAGGCTCCTGGGGACCC 1264  
Qy 179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198  
Db 1265 GTGTGTGATGATAGTGGACACCACTGACGCAATGTGTGTGGCTGGCGGAGCTGGCTGT 1324  
Qy 199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218  
Db 1325 GGCTGGCCACGTCAGCCCGCAATGCGCGTTGGCCAGGTTTCAGGACCCATTTGTC 1384  
Qy 219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238  
Db 1385 CTGGATGACGTGCGTGTCTGCTCAGGCTATGAGTCTTACCTGTGGAGCTGCCCGCACAAATG 1444  
Qy 239 TrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTrp----- 255  
Db 1445 TGGCTCTCCCAATACTGTCAGCACAGTGAAGACGCTGGTGTCTGCTCAGCTGCCAC 1504  
Qy 255 ----- 255  
Db 1505 TCTGGTGCAGCCCGCAGTCCAGACAGTGGCCGACCATCACCATTACCTGTCATCGACAGTA 1564  
Qy 256 ---AspSerSerAspLeuGluLeuArgLeuValGlyThrAsnArgCysMetGlyArg 274  
Db 1565 GGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGAGGTGACAGGTGTCCAGGCGCA 1624  
Qy 275 ValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAla 294  
Db 1625 GTGGAGGTCTCTACCGAGGCTCTGGGCGACCGTGTGTGATGACAGTGGGACACCAAT 1684  
Qy 295 AlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeu 314  
Db 1685 GATGCCAATGTGCTGCGAGGACGCTGGCTGTGGCTGGCCATGTGGCCCCCAGGAAT 1744  
Qy 315 ProHisLeuGlnSerGlyAspValValTrpLeuAspGlyValSerCysSerGlyAsn 334  
Db 1745 GCCCGGTTTGGTCAGGCTCAGGACCACTATGTCTGGATGACGTGCGCTGCTCAGGGAAT 1804  
Qy 335 GluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGln 354

Db 1805 GAGTCTCTACTTGTGGAGCTGCCCGCACAAATGGCTGGCTCTCCCAATACTGGGCCATAGT 1864  
Qy 355 AsnAspValSerValIleCysSer---AspGlyAlaAspLeuGluLeuArgLeuAlaAsp 373  
Db 1865 GAAGACGCTGGTGTCTATCTGCTCAGGACCTGAATCCAGTTTGGCCCTGAGGCTGTGTAAT 1924  
Qy 374 GlySerAsnAsnCysSerGlyArgValGluValArgIleHisGlnGlnTrpTrpThrIle 393  
Db 1925 GGAGGTGACAGGTGTTCAGGCGCGAGTGCAGTCTATATACCGAGGCTCTTGGGCGCACCGTG 1984  
Qy 394 CysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysPro 413  
Db 1985 TGTGATGACAGCTGGGACCAACCAATGATGCCAATGTGGTCTCAGGACCTGGGCTGGC 2044  
Qy 414 PheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaAspIleTrpIle 433  
Db 2045 TGGGCCATGTGAGCCCGAGGAATGCCCGTTTGGTCAGGCTCAGGACCCATTTCTCTG 2104  
Qy 434 AsnSerIleSerCysThrGlyAsnGlnSerAlaLeuTrpAspCysThrTrpAspGlyLys 453  
Db 2105 GATGATGTGCTGCTCAGGACATGAGTCTTACCTGTGGAGTGCCTCCCAACAATGGCTGG 2164  
Qy 454 AlaLysArgThrCysPheArgSerAspAlaGlyValIleCysSer----- 469  
Db 2165 CTCTCCCAACTGTGGCCATCATGAAGATGCTGGTGTCTGCTCAGTGCCTGCCAGTCC 2224  
Qy 469 ----- 469  
Db 2225 CGTCCGCGCCCGGCGCAGACAGCTGTGCGACCATCAGCTTACCTCCATCGACAGTAGGA 2284  
Qy 470 AspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTrpGlyArgLeu 489  
Db 2285 TCTGAATCCAGTTTACCTGAGGCTGGTGAATGAAGTGACAGGTGTGAGGCGCGAGTA 2344  
Qy 490 GluValLysTrpGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsn 509  
Db 2345 GAGTGTCTATACCGAGGCTCTGGGCGACCGTGTGTGATGACAGCTGGATACCAATGAT 2404  
Qy 510 AlaAlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThr 529  
Db 2405 GCCAATGTGTGCGAGGACGCTGGCTGGTGGCCATGTCAGCCCGCAGGAAATGCC 2464  
Qy 530 TyrPheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGlu 549  
Db 2465 CGGTTTGGCCAGGCTCAGGACCCATTTCTGTGATGATGTGCGCTGCTCAGGACACGAG 2524  
Qy 550 SerAsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGlu 569  
Db 2525 TCTTACCTGTGGAGTGTCCCGCACAAATGGCTGGCTCTCCACAACTGTGGCCATCATGNA 2584  
Qy 570 AspValIleValThrCysSer-----GlyAspAlaThrTrp 581  
Db 2585 GATGCTGGTGTCTATCTGCTCAGTTCCTCCAGTCCCGCGCGACACCCAGTCCAGATACTGG 2644  
Qy 582 -----GlyLeuArgLeuValGly 587  
Db 2645 CGGACCTCACATGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGTGTAAT 2704  
Qy 588 GlySerAsnArgCysSerGlyArgLeuGluValTrpPheGlnGlyArgTrpGlyThrVal 607  
Db 2705 GGAGTGCAGAGGTGTGAGGCGGAGTGGAGGTCCTATACCGAGGCTCTGGGCGCACCGTG 2764  
Qy 608 CysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysPro 627  
Db 2765 TGTGATGATAGTGGGACACACAGTGCAGCCCAATGTGGTGTGGCGGAGTGGCTGTGGC 2824  
Qy 628 SerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTrpGlyLysIleTrpLeu 647  
Db 2825 TGGGCGAGTGCAGCCCGCAGGAAATGCCCGTTTGGCCAGGTTTCAGGACCCATTTGCTCT 2884  
Qy 648 AspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrp 667

Db 2885 GATGACGTGGCTGCTCAGCGTATGAGTCTTACCTGTGGAGCTGCCCCCAATGGCTGG 2944  
Qy 668 GlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerAspAla----- 685  
Db 2945 CTCTCCCATTAACGTGACGACAGTGAAGCGTGGTGTCTATCTGCTCAGCTGCCCACTCC 3004  
Qy 685 ----- 685  
Db 3005 TGGTCGACGCCAGTCCAGACACATTGCCGACCATCACTTGGCTGCTGCATCGACAGTAGGA 3064  
Qy 686 -----SerAspMetGluLeuArgLeuValGlySerSerArgCysAlaGlyLysVal 703  
Db 3065 TCTGAATTAAGTTGGCCCTTGAGCGTGGTGAATGGAGGTGACAGGTGTCAGGCGCGAGTG 3124  
Qy 704 GluValAsnValGlnGlnAlaValAlaGlyLeuLeuCysAlaAsnGlyTrpGlyMetAsnIle 723  
Db 3125 GAGGTCTATACCAAGCTCTCTGGGACCGTGTGGCATGACAGCTGGGACACCAATGAT 3184  
Qy 724 AlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluPro 743  
Db 3185 GCCAATGCTGTCAGGCAACCGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3244  
Qy 744 HisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlu 763  
Db 3245 CGGTTTGGTCAGGCTCAGGACCATTTGCTCTGGATGATGCTGCTCAGGACACGAG 3304  
Qy 764 AlaSerLeuTrpAspCysIleArgTrpGluTrpIleGlnThrAlaCysHisLeuAsnMet 783  
Db 3305 TCTTACCGTGGAGCTGCCCCCAATGGCTGGCTCTCCCACTGTGGCCATAGTGAA 3364  
Qy 784 GluAlaSerLeuIleCysSerAlaHisArg---GlnPro----- 795  
Db 3365 GAGCGTGTGTCATCTGCTCAGCTTCCAGTCCCGGCAACACCTAGTCCAGACACTTGG 3424  
Qy 796 -----ArgLeuValGly 799  
Db 3425 CCAACCTCAGATGCATCAACAGCAGGAGTCTGAATCCAGTTTGGCCCTGAGGCTGGTGAAT 3484  
Qy 800 AlaAspMetProCysSerGlyArgValGluValIleHisAlaAspThrTrpArgSerVal 819  
Db 3485 GGAGGTACAGGTGTACAGGCGGAGTGGAGGTCTATACCGAGGCTCTCTGGGCGACCGTG 3544  
Qy 820 CysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGly 839  
Db 3545 TGTGATCACTACTGGGACACCAATGATGCCAATGTGTTGTCAGGACGCTGGCTGTGGC 3604  
Qy 840 AspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAla 859  
Db 3605 TGGGCCATGTACGCCCCAGGAAATGCCCGTGGTGGCCAGGTTTCAGGACCCATTTGCTCTG 3664  
Qy 860 GluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHis 879  
Db 3665 GATGATGGCGTGTCTCAGGACATGATGCTCTATCTGTGGAGTGCCTCCCAATGGCTGG 3724  
Qy 880 ProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAsp 899  
Db 3725 CTCTCCCAACAATGTGGCCATCATGAAGACGCTGGTGTCTGCTCAGCTTCCAGTCC 3784  
Qy 899 ----- 899  
Db 3785 CAGCCGACACCCAGCCAGACACTTGGCCAACTCAGATGCATCAACAGCAGGATCTGAA 3844  
Qy 900 -----ValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIle 915  
Db 3845 TCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTACGGGCGGAGTGGAGGTC 3904  
Qy 916 AsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArg 935  
Db 3905 CTATACCGAGGCTCTGGGCGACCGTGTGTGATGACTACTGGGACACCAATGATGCCAAT 3964  
Qy 936 ValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIle 955  
Db 3965 GTGGTTGACAGGAGCTGGGCTGTGGTGGGCGCACGTCAGCCCCAGGAAATGCCCGGTTT 4024

Qy 956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975  
Db 4025 GCCCAGGGTTCAGGACCCATTTGCTCTGGATGATGTGCTGCTCAGGACATGAGTCTCTAT 4084  
Qy 976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrVal 995  
Db 4085 CTGTGGAGCTGCCCCCAATGGCTGCTCCCACTGCTGGCCATCATGAAGACGCT 4144  
Qy 996 SerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSer 1015  
Db 4145 GGTGTCATGCTCAGCTTCCAGTCCAGCCGACA---CCAGCCCAACACTTGGGCCA 4201  
Qy 1016 AspProTrpLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArg 1035  
Db 4202 ACCTCAGATGCATCAACAGCAGGATCTGAATCCAGTTG-----GCC 4243  
Qy 1036 LeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGly 1055  
Db 4244 CTGAGGCTGTGAATGGAGGTGACAGGTGTCAAGGCGGAGTGGAGTCTATATCCGAGGC 4303  
Qy 1056 PheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGln 1075  
Db 4304 TCTGGGCGACCGTGTGTGATGACTACTGGGACCACTGATGCAATGTGGTTGAGG 4363  
Qy 1076 LysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySer 1095  
Db 4364 CAGCTGGCTGTGCTGGGCGACCTCAGCCCGAGGAAATGCCGTTTGGCCAGGTTCA 4423  
Qy 1096 GlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCys 1115  
Db 4424 GGACCCATTTGCTGGATGATGTGCTGCTCAGGACATGATGCTCTATCTGAGGCTGC 4483  
Qy 1116 ProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCys 1135  
Db 4484 CCCCAATGGCTGGCTCTCCCACTGTGGCCATCATGAAGACGCTGGTGTCTATCTGC 4543  
Qy 1135 ----- 1135  
Db 4544 TCAGCTTCCAGTCCCGGACACCCAGCCAGACACTTGGCCCAACCTCTCGTGCATCA 4603  
Qy 1136 -----SerGluPheThr---AlaLeuArgLeuTrpSerGluThrGluThrGluSer 1151  
Db 4604 ACAGCAGGATCTGAATCCACTTTGGCCCTGAGACTGTGTAAT-----GGAGGTGACAGG 4657  
Qy 1152 CysAlaGlyArgLeuGluValPheTrpAsnGlyThrTrpGlySerValGlyArgArgAsn 1171  
Db 4658 TGTGAGGCGGAGTGGAGGTCTTATACCAAGGCTCCTGGGCGACCGTGTGTGATGACTAC 4717  
Qy 1172 IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal 1191  
Db 4718 TGGGACACCAATGATGCCAAGCTGTGTCAGGACGCTGGGCTGT-----GGCTGG 4768  
Qy 1192 ValSerLeuAlaProLeuSerLys-----ThrGlySerGlyPheMetTrpValAsp 1208  
Db 4769 GCCATGTAGCCCCAGGAAATGCCAGTTTGGCCAGGCTCAGGACCACTTGTCTGAT 4828  
Qy 1209 AspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp--- 1227  
Db 4829 GATGTGGCTGTCTCAGGACACGAGTCTTACCTGTGGAGCTGCCGCCCAATGGCTGGCTC 4889  
Qy 1228 -----GluArgArgIle 1231  
Db 4889 TCCCACTGCTGGCCATCATGAAGATGCTGTGTCATCTGCTCAGTGTCTCAGTCCAG 4948  
Qy 1232 SerSerProAlaGluGluThrTrpIleThr-----Cys 1242  
Db 4949 TCAAGGCCAGGCCAGATCTTGGCTGACCACTTACCGGCATTCAGAGTAGGATCT 5008  
Qy 1243 GluAspArgIleArgValArg-----GlyGlyAspThrGluCysSerGlyArgVal 1259  
Db 5009 GAATCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGAC---AGGTGTGCGAGGCCGAGTG 5065

28. CysLeuLeuLeu-----AsnSerCysPheLeuIleSerSerPheAsnGlyThr 44



Db	79914	TGATGGCATCCTGGCAAGCCATTTTCAGGCTCAGTGAAGGCTCATGAAGTGCACAGTG	79855
Qy	458	-----CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAsp	475
Db	79854	TCCTGCTGCTGGCAGGAAGTGGGGTGTCTATCAGTCATAGTACCCACACAGGTGGCT	79795
Qy	476	LeuArgLeu-----ValGlyAlaHisSerProCysTyr	486
Db	79794	CTCAGGCTCTGGGAGATGCTTCAGCTCTCTTTATCATGGGGAACACTCC-----	79744
Qy	487	GlyArgLeuGluValLysTyrGln-----GlyGluTyr	497
Db	79743	-----TCTGTGAACATCCCATTTCCAGTACATCAGACACTGTGTGAATAGATGG	79693
Qy	498	-----GlyThrValCysHis-----	502
Db	79692	CTGGGACCTAGACACACTGCCAGTCCAGCCAGCATGGCACCAGCTGCAGCCCTCCAGATG	79633
Qy	503	-----AspArgTyrSerThrArgAsnAlaIleVal-----	512
Db	79632	GATAAGGAGGTTATCAGTGGGGCTGTAGACATGTAGAGCTTAGGGCTTACTGGGGCCC	79573
Qy	513	-----ValCysLysGlnLeuGlyCysGly-----	520
Db	79572	AGGACAGGATGCAATCTTGTGGGAGCTGGACTCTCAAAATGGTGCCACTGCTGGAGTTACC	79513
Qy	520	-----	520
Db	79512	TGGTCTTGGGGTTGGGATTTGGATGTAGGACCCAGCAAAACTCCCCCTCTGGAAACA	79453
Qy	521	-----LysProMetHisValPhe	526
Db	79452	CGCTATTGTGCAAACTCTAGGCAGCTCCCCTGTACTAGTCTCAGAGCCCAAGAGCTGAG	79393
Qy	527	GlyMetThrTyrPheLysGluAlaSerGlyProIleTyr-----	539
Db	79392	GGGCTCTCCCATGGCTAGGATTGCAAGGCCCAAGTGGGAATCTGGACTCTAGGAATC	79333
Qy	539	-----	539
Db	79332	TCTCATTTATCTTTTCTACAGTGGGGATCCATACCAGGCTTCCAGTCAATCCTAGCCA	79273
Qy	539	-----	539
Db	79272	GACAATTGCTTCACCTCTTCTCTCTGTGTGCGCTCAGAAGATCTGTTCCTATCACTTC	79213
Qy	540	-----LeuAspValSerCys-----	545
Db	79212	CCTGCTGGATTCCAGTGTCTCTATCTTAGACACTATTCTGTTGTGAATATATCTACTTCTG	79153
Qy	545	-----	545
Db	79152	TTTTGATCGTTTTTGGTGGTGAAGTGTGGGTGCGCTTAGTAGCGGCTTGAAGCGCTCCCC	79093
Qy	545	-----	545
Db	79092	CTTGAGCTATTCTCTTTAAGGGAAAGTTCCTCTCTCCATAATCATGAATATCAATGTGT	79033
Qy	545	-----	545
Db	79032	CTTTCAATAAGAGAGTTGAGGATTCCTGAACACGGGACCAACTCCAATCCATAGAAACT	78973
Qy	546	-----IleGlu	547
Db	78972	ACCACACACCTTCTCTCTGTCTGTGATCTGTATATGTAAACTACAGACTTCTCTATTGG	78913
Qy	547	YasnGlu-----	549
Db	78912	TAATAAATGTTTTACATATTGAGAAATTATATATATATCTCTATGTGTATATATATT	78853
Qy	549	-----	549

Ds 78852 TAAATAAATTATATATAAACAATTATATATAAACAATTCTATATAAACCATTATAT 78793  
Qy 549 ----- 549  
Ds 78792 ATAAATATATAATTATATATAAACAATTATATATAAATTATATATATATATATAT 78733  
Qy 549 ----- 549  
Ds 78732 AAAAACCTAAAAGTCAATATATGATGTCACCTTTTATGCTGAATCAGCTACAGCAACTGT 78673  
Qy 549 ----- 549  
Ds 78672 TTTCCACCTAAGACTCTCTCTCAGAGTGAATATGTTCTACAAAGCAGTGAATATATGTT 78613  
Qy 550 -----SerAsnIleTTPAspCysGluHisSerGlyTTPGlyLysHis-- 563  
Ds 78612 TTTCCCTTATATTTAGAGTAGTACTTATTAATGAAGAAGTTTCTGTTCTAGACACCT 78553  
Qy 563 ----- 563  
Ds 78552 TTTTCAGGAATCTCAATTTTGATAGAGATTCCTACAGATATATATGTTTATTTCTCCCA 78493  
Qy 563 ----- 563  
Ds 78492 CCCTCCCAACCAACACACACAAAAATTATAGCTCTCTAAACGTCAGATCTAAAAATAAAC 78433  
Qy 564 -----AsnCysValHisArgG1 569  
Ds 78432 CCATTTAAAAAATTCTTTTATGTTTATGTTAAAGTGTGAAGCTGATCTACCAACC 78373  
Qy 569 uAspValIleValThrCysSerGlyAspAlaThrTTPGlyLeuArgLeuValGlyGlyse 589  
Ds 78372 TCATGATGATTTGTT-----TTGAACCTCCTGCCAATTAC 78337  
Qy 589 rAsnArg-----Cys----- 592  
Ds 78336 AAACAGATCTTGACTCAATGTTTTCACAAACCTGTCAGCTGAGCTCCAAAGACACAGATA 78277  
Qy 593 -----SerGlyArgLeuGluValTyrPheGlnGlyArgTTP----- 604  
Ds 78276 CAATAGGCTTAAATGGAATACATTAACAACCTCTATTTCTGCTGAATGGAGTTTAA 78217  
Qy 604 ----- 604  
Ds 78216 TTTAGAGATCATCTCATATCTACACAGATTTATATCAAAATAACCAATGATGATAAAT 78157  
Qy 605 -----GlyThrValCysAspAspGlyTTP----- 612  
Ds 78156 TCATTCAGCTGTTTATAAGTGGCAGGAAGTGGAGCTGA-AGTGATAGGAAGGACAGAC 78098  
Qy 612 ----- 612  
Ds 78097 TTTCCAGGCTCTAGACCAGTTAATATATATATTTCTTCTTCACAGGCCAATAGGAAG 78038  
Qy 612 ----- 612  
Ds 78037 TAACTTGCTCAGTTTGATCTGGTCAAGTCAGAGTTGATTTTCTGACATGTCACATCTG 77978  
Qy 612 ----- 612  
Ds 77977 TCAACATTAGCAGTTAATAAACCTTATTAATATAATTTAGTATATAAAGTAGAATAATAG 77918  
Qy 613 -----AsnSerLy 615  
Ds 77917 TATGATATCAGATAGACATTTAATAATGTGAATTTTATTTTAACTAAACAATAGAAA 77858  
Qy 615 sAlaAlaAlaValValCysSer-----GlnLeuAspCys----- 626  
Ds 77857 AGAGCTGATACCATTTATAGTTTAAATAGATAAAGAAGTATTCATTTGTATATATAT 77798  
Qy 626 ----- 626  
Ds 77797 TAGATATATTTCTCAAAATTAGTAGCAATATTTATTTGTTTTCACAGACTATGCCACAGCA 77738

Qy 626 ----- 626  
Ds 77737 CACTCAAAAGCTACCAAACTGCCCTAAAAGATAAAGTAACCTTTCTAGAAATGAGATTCAT 77678  
Qy 626 ----- 626  
Ds 77677 TTTCAATTTTATGTCATCTGCTGCTAAAAAATTTCTTTCAAAAGCTGACTCTTTCAATAGGCT 77618  
Qy 626 ----- 626  
Ds 77617 CAGAATTAATTTCTGATTTAATTTCTGCTTTCTCTCTTTTAAAAAAGTTTTTATTTTCCA 77558  
Qy 626 ----- 626  
Ds 77557 TAGGTTTTTAGGGAACAGATGGTAATAGTTACATGAGTAAGTTCTTTAGTTTGTGTTGT 77498  
Qy 627 -----ProSerSer----- 629  
Ds 77497 GAGATTTTGATGCACTCATCACCAAGCAGTATACACTGAACCCAATTTGTAGTCTTTTC 77438  
Qy 630 -----IleIleGlyMe 633  
Ds 77437 TTTCTCAAAACCTCGTACCTTTTCCCTGAGTCTCTCAAAGTCCCATTTGATTTCTTAT 77378  
Qy 633 tGlyLeuGlyAsnAlaSerThrGly----- 641  
Ds 77377 CTCCTTGATCGTCATAGCTTTAGCCCCCCTTATGAGTGAGAACATACATATTTGGTT 77318  
Qy 641 ----- 641  
Ds 77317 TTTCCATTTCTGATTTACTTTACTTTAGTAATAATAGTCTTCAATTTCCATCTGTTGCTGTG 77258  
Qy 642 -----TyrGlyLysIleTTPLeuAspVal---SerCys----- 652  
Ds 77257 AATGCCATTAATTCATCTCTTTATGCTGAGTAGTATTCATCATGCTGTATATATATA 77198  
Qy 652 ----- 652  
Ds 77197 CACACATACATGTGTGTATACGCATATACACATACATGTATGTATACGGGTATACACA 77138  
Qy 652 ----- 652  
Ds 77137 CATACATGTATATACGTGTATACATATATATATATATATATATGTATGTATATACACACA 77078  
Qy 653 -----AspGly----- 654  
Ds 77077 CACACACACACACACACACAGTTTCTTTATCCACTTTGTTGATTTGAGGCAATTTGGG 77018  
Qy 654 ----- 654  
Ds 77017 CTGGTTCCATATTTTGCATTTGCGATTATCTCTCTATATAACATGCGTGTGCAAGTATC 76958  
Qy 655 -----AspGluSerAspLeuTTPSer----- 661  
Ds 76957 TTTTTCAGATAATGACTTATTTTCTCTCTGGGTAGATACCTAGTAGTGGGATTTGCTGATC 76898  
Qy 661 ----- 661  
Ds 76897 AAATGGTAGTCTGCTTTTATGTTTCTTTAAAGAAATCTCCACTGTTTTCATACGCGTTGTA 76838  
Qy 662 -----CysArgAsn----- 664  
Ds 76837 CTAGTTACATTTCCCAACAGCAGTGTAGAAAGTGTCCCTTTCCACTGCATCCATGAAACT 76778  
Qy 665 -----SerGly----- 666  
Ds 76777 CTATTTCTTTTGTATTTTGTGTTATGCCATTTCTTTCGAGGAGTAAAGTGGCATTTGCATT 76718  
Qy 666 ----- 666  
Ds 76717 GTGGTTTTGATTTGATTTTCTCTGATCATTTAGTATGATGAGCAATTTTTCATATAAAAAA 76658



Db 74498 GACATCAATTCCAAAAAGCCCTAAGAGCCCAATATATAGATGAACACTGAACAAATTTCCACT 74439  
QY 790 ----- 790  
Db 74438 AAGCTTAGCTACAGAGAGTCATGAATAAATTTAATAAAGTAGTTATGTGCACATTAGAA 74379  
QY 790 ----- 790  
Db 74378 AAAGGAGATCATTTCAAGTTGAGAAAGAAATGGGAGGTAAAGATGTTGGGATATAGGCA 74319  
QY 790 ----- 790  
Db 74318 TAGACTATTTCCAAAAAGCTGTATTTTAAAGGGGAAGAAATAAATAAAGTGATGACTAAAA 74259  
QY 790 ----- 790  
Db 74258 AGGAATATAATCTTCAGGGGAGTTACTCAATTTTTTTTATTTGATAGGAGATCTAGAGTA 74199  
QY 790 ----- 790  
Db 74198 GTGCTGTTCAAAAGAACATCACCTTCTGTGTAGTAAGTAATAATATATGCTATTAGCTGATG 74139  
QY 790 ----- 790  
Db 74138 TTAGAGGCTATAGGAAAAATAGAACTAGAAATAGAAATCCAGAGTGTTTATGGGACAGA 74079  
QY 790 ----- 790  
Db 74078 ATGAATTTCCAAATAAATGACTGGGTAGGCCCTTACTGAAAAGGTGAAACATAGACAA 74019  
QY 790 ----- 790  
Db 74018 GACTCGAAGGTGATGGGAGAGTCAGTTATATGTCATATACGAGAAGGAAAAATACAAATAG 73959  
QY 790 ----- 790  
Db 73958 AGATACAGCCAGTTCAAAGCCTTGGGGCAAGATTATGCCCTGATATGTTTGAAGAAGAGA 73899  
QY 790 ----- 790  
Db 73898 AAGGAATCCATGCAAGTGAAGAGCAGTAACTAAACATCAGGACCTGGCACCAGG 73839  
QY 790 ----- 790  
Db 73838 GTAAATGAGCAGGGGAAGGTCACTCAGGGCAAGAGTCTTCGTAGGAGTTTAGATTTT 73779  
QY 790 ----- 790  
Db 73778 TACTCTGAGTGAATATGGAGTCATGGGAGGATTTGGGGCAGAATACTGATTTAACTGA 73719  
QY 790 ----- 790  
Db 73718 TACATCATTTTAAATAACCTCTCTGGATTTTATTTTTTAAAGATGAACATATAGAGGA 73659  
QY 790 ----- 790  
Db 73658 ATAATTATAGAAACAGGGAATCTCTAGGAATCTATTATAGTAAGTCACTCAGGCAAGTAATGAT 73599  
QY 790 ----- 790  
Db 73598 GGTGGCTCAATCAAGGCCCTAGGGTTGAAGGGGTGAGAAGTTTGGGATTTATGGGCATAC 73539  
QY 790 ----- 790  
Db 73538 TTTGAAATACAGAAACAAACTTTCTGAAGGGGAGATATAGTATGTAATGGAAGTCA 73479  
QY 790 ----- 790  
Db 73478 AGAATTATTTCAAAATATTTGAATGCTCTGGAATGATGGAGTTCATCAATAGAAATATGG 73419  
QY 790 ----- 790  
Db 73418 AAGTCTGAGCCTTAGATAGATGTTTGAAGAGGACTCCAAAGAAATTCATTTAGATATCT 73359

QY 790 ----- 790  
Db 73358 ATAAGAAATCCAAATGGAATGCTAAGTAGAAATTTGGATAAATGAGTCAGGAGTTTGAA 73299  
QY 790 ----- 790  
Db 73298 AAATGCCTGGTTGGAGACAGATAACATACATTACTAATAAAGTAGCTTTGCTGCAAG 73239  
QY 790 ----- 790  
Db 73238 AATAGAGTCAGAAATAATGGGAGTATTTTTTGTGTTTATAAATTAATCTAATTTAGATGG 73179  
QY 790 ----- 790  
Db 73178 ATTGCTGGATCTAACCAAGTAGAAGTCTTTCTATTCTTATATTATACACATCTTCTT 73119  
QY 790 ----- 790  
Db 73118 AAGATGACTACAAATCTTCACGTGTTGTTAATGCTGAGGTTACTTAGTATTTTTTAAT 73059  
QY 790 ----- 790  
Db 73058 GTAACAAATTTGTCCAATTCATTTCTTCGATGAGTTCTAAATCTAGGATTTGGTGGTC 72999  
QY 790 ----- 790  
Db 72998 ACCAAGACATAGAAATAGGTATGTAATCCAGATCTGTGAATTTAAGTCTGACTGCTAA 72939  
QY 790 ----- 790  
Db 72938 TCTTGTCTCTGAGCCCCCTACGAATGGGATGTTCTTAATGATGAATGGAAAAATGCTC 72879  
QY 790 ----- 790  
Db 72878 TGGGTTAGATATGGAGACTCGAGACAGGAGATAGATATTGGAACCCCTAAACATAACC 72819  
QY 791 -----AlaHisArgGlnProArgLeuValGlyAlaAspMetProCy 804  
|||||  
Db 72818 TTTGAAATATTTATTTATAGCCACAGCCAGCGCTGTTGGAGCTGATATGCCCTG 72759  
QY 804 sSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspH 824  
|||||  
Db 72758 CTCTGGAGCTGTTGAAGTGAACATGCAGACATGCGCCTCTGCTGTGATTTCTGATTT 72699  
QY 824 eSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLe 844  
|||||  
Db 72698 CTCCTTCATGCTGCCAATGCTGTGCAGAGAATTAACCTGTGGAGATGCCATATCTCT 72639  
QY 844 uSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCy 864  
|||||  
Db 72638 TTTCTGTGGAGATCACTTTGAAAAAGGAATGGTCAACTTTGGCGGAAAAAGTTCCACTG 72579  
QY 864 sGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCy 884  
|||||  
Db 72578 TGAAGGGAGTGAACATCACCTTGCAATTTATGCCCTTTCAACATCCGGAAGACACTG 72519  
QY 884 sIleHisSerArgGluValGlyValValCysSerArg 896  
|||||  
Db 72518 TATCCACAGCAGAGAAAGTTGGAGTTGCTGTCTCCCG-TGAGATTTTAAATAATCTTAAT 72460  
QY 896 ----- 896  
Db 72459 TGGCTGGGCTAGAGATGGGAGTTTGGGCTGGAGCTGATAAGTCATGATGTCCTCTCT 72400  
QY 897 -----TyrThrAspValArgLeuValAsnGlyLysSerGlnCy 909  
|||||  
Db 72399 ATGCCCTTGTCTCTCTCTAGGATATACAGATGCTCCGACTTGTGAATGGCAATCCCACTG 72340  
QY 909 sAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTr 929  
|||||  
Db 72339 TGACGGGCAAGTGGAGATCAACGCTGTTGGACACTGGGGCTCACTGTGTGACACCCACTG 72280

QY 929 pAspProGluAspAlaAArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerTh 949  
|||||  
Db 72279 GGACCCAGAGATGCCGCTGTTCTATGCAGACAGCTCAGCTGGGACCTCTCTCAAC 72220  
QY 949 rThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCy 969  
|||||  
Db 72219 CACAGAGAGAAATATATTGGAGAAAGAGTGTGCTGTGGGGACACAGGTTTCATG 72160  
QY 969 sLeuGlyAsnGluSerLeuLeuAspAsnGlnMetThrValLeuGlyValAlaProProCy 989  
|||||  
Db 72159 CITAGGGAATGAGTCACTTCTCGATACCTGTCAATGACAGCTTCTTGGAGCACCCTCCG 72100  
QY 989 sLeuHisGlyAsnThrValSerValIleCysThr 1000  
|||||  
Db 72099 TATCCATGGAATACTGTCTGTGTATCTGCACAGGTAAGAGAATAGTCTTATGGTTAA 72040  
QY 1000 1000  
Db 72039 CTACTGTCAATATAGCCAGACATGTTAGAGGAGTTGGGATTAAGAGAAAGATGAGTAA 71980  
QY 1000 1000  
Db 71979 GAAATGGGGTGCAGTGGAGGGGATAGGCATCACAGTCAAGAGTAACATGTAAGA 71920  
QY 1000 1000  
Db 71919 AATAATGATACTAGGACCATCTTAAGGGGAAGTTAACACCTCTGTGTGGGATTTAC 71860  
QY 1000 1000  
Db 71859 GAGTAGACATCATTAATAAATGATTCAATAGAGCAGGCTTGAAGAGTCTTATC 71800  
QY 1000 1000  
Db 71799 ACAATGTCAGATCCATGTTGTACAGAACTCTTCAACCTTTTCCATGATTTTGGTCT 71740  
QY 1001 1010  
Db 71739 ACTTTGGCCCATTTCTGCACCCCTCCACAGGAAGCCCTGACCCAGCCACTGTTTCCATG 71680  
QY 1010 sLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCy 1030  
|||||  
Db 71679 CCTGCAATGATCTGACCCATATTTGCTGCAGTCCAGAGGCGAGTGTGATCTG 71620  
QY 1030 sLeu 1031  
Db 71619 CTT-AGGTAAGCAGAATCAGTTTTAGGAAACACTGCAACTCTAAAGAAATGTGACCTATTA 71561  
QY 1032 1045  
Db 71560 ACTGACCTTCTCCTCTAAGAGGACAAAGCGCTCCGCTAGTGGATGGGACAGCCGCTG 71501  
QY 1045 sAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspGlyTr 1065  
|||||  
Db 71500 TGCCGGAGAGTAGATCTATCAGACGGCTTCTGGGCAACCTCTGTGATGACGGCTG 71441  
QY 1065 pAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAl 1085  
|||||  
Db 71440 GGACCTGAGCGATGCCACGTGGTGTGTCAAAAGCTGGGCTGTGAGTGGCCCTTCAATGC 71381  
QY 1085 aThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspLeuAsnCy 1105  
|||||  
Db 71380 CACGGTCTCTCACTTTGGGGAGGGGTACAGGCCCACTGCTGGATGACCTGAACCTG 71321  
QY 1105 sThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCy 1125  
|||||  
Db 71320 CACAGGAATGGAGTCCCACTTGTGCGAGTGCCTTCCCGCGCTGGGGCAGCAGCTG 71261  
QY 1125 sArgHisLysGluAspAlaGlyValIleCysSer 1136  
|||||  
Db 71260 CAGGCACAAGAGGAGCGAGGGGTCTCTGCTCAGGCTGTGCTGTCTACCTTTTCCACCT 71201  
QY 1136 1136

Db 71200 CCTCTACACACACAGTGCACACACTTGGCACACACACACAATTTGTAAGTCTAGTGAAA 71141  
QY 1136 1136  
Db 71140 AGTGTATGGCCTTAGGAATACAAAACAAAATCCACTTGGTTTTTAATCCTATTCTCTATA 71081  
QY 1136 1136  
Db 71080 CTGATTTCTTAATACCCCTTGTCTATATTACATTTGTCTGGGTATCATTCATAACTGAGAAG 71021  
QY 1136 1136  
Db 71020 AATAAAAGATTGGTCTTACTGTCTAGAGAATGTTAGACCCCAAGAAAGGACCTTACATCA 70961  
QY 1136 1136  
Db 70960 CATCOACTCTGGCTCTGAGAACACACTGTCAACCCGTGTCCAATCTGCCCTATTGTTGT 70901  
QY 1136 1136  
Db 70900 TTTGTTTCTTTTATTTTAGGCTTTTAGTGCCTGAAGCCATGGTTATTAGTTCTGTCT 70841  
QY 1136 1136  
Db 70840 CTAGTGTGAAGAGAAAGAGGATGAGGAAGGCTCTTACTAACCTGACAGAAAGCAGAA 70781  
QY 1136 1136  
Db 70780 ACTAAGAACCCATACTGTATCTCTCCCTTGGACACCCCTGCAANTGAGGTTAACCTCC 70721  
QY 1136 1136  
Db 70720 TGATTTTCTGTAATTTGAATTTATGTTTCGAGAGTAGAATGTAGAAGGTGTCATTTAAG 70661  
QY 1136 1136  
Db 70660 TTAAGTTGATACAAATTTAAAGCTCAAATACTATTCTTTGGGAAACCCACACTCATCTGA 70601  
QY 1136 1136  
Db 70600 ANTCTACAAGACCCCTTAGTTTCCATCAGATTTATGCTGTGTTAATATGTTAGAGTATTA 70541  
QY 1136 1136  
Db 70540 TGAATTTGTTGTGAGCCCTAATGTATTATTTCATGGGAAAAATGGATTCCTTTGGGA 70481  
QY 1137 1138  
Db 70480 GGGATTTCCAACTCCAAGGAAATATTTAATACTAAGTTTTTGTATTCTCTGTAGAATT 70421  
QY 1138 eThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluVa 1158  
|||||  
Db 70420 CACACCTTGGGCTCTACAGTGAACCTGAACACAGAGAGTGTGCTGGGAGATTGGAAGT 70361  
QY 1158 lPheTyrAsnGlyThrTrpGlySerValGlyArgAsnIleThrAlaIleAla6I 1178  
|||||  
Db 70360 CTTCTATAACGGGACCTGGGCGAGCTCGGAGGAGAACATCACACAGCCATAGCAGG 70301  
QY 1178 yIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSe 1198  
|||||  
Db 70300 CATTTGTGAGGAGCTGGGCTGTGGGAGAAATGAGTTGTCAGCTCGGCCCTTTATC 70241  
QY 1198 rLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSe 1218  
|||||  
Db 70240 TAAGACAGCTCTGGTTTCATGTGGGTGATCAGTTCAGTGTCTTAACAGCATATCTC 70181  
QY 1218 rIleTrpGlnCysLeuSerAlaProTrpGluArgIleSerSerProAlaGluGluTh 1238  
|||||  
Db 70180 CATATGGCAGTGCCTGTCTGCCCATGGGAGCAAGAAATCTCCAGCCCGCAGAGAGAC 70121  
QY 1238 rTrpIleThrCysGlu 1243  
|||||

Db 70120 CTGGATCACATGTGAAGTGGGTATCTTTCTAAACACAGCTGTCACTGAAGCCCTGTGA 70061  
QY 1243 ----- 1243  
Db 70060 AATATAACTTACCATAAATAGCAAAATCTTCCATTACAGGACACATGCGAGACTCACTGA 70001  
QY 1243 ----- 1243  
Db 70000 AAAGTGCATGACAATAAATGCTGATTTATCGCATCTGTCTATATTGAAATTTTATTATT 69941  
QY 1243 ----- 1243  
Db 69940 TTGGTTTAGTATTTTACAGGGTATTTAAATAATGTATATGTCTCTTTTGAATCCTTTTAAG 69881  
QY 1243 ----- 1243  
Db 69880 CTCCTTGTGAGAGCTGGATGAAGCACATANAAGGATAACACACATANAANAATTACAG 69821  
QY 1243 ----- 1243  
Db 69820 TGTTTTAAAAAATGAAAAATATAATATGTACTGTCTTAGAAATCTAGAAAGCTCAGA 69761  
QY 1243 ----- 1243  
Db 69760 ACTTTTGAGGAATGTGAATTAATFATGGAGATTAAAGCAGCTCTATTTTTCCTCCAGAA 69701  
QY 1243 ----- 1243  
Db 69700 CATTTTACTATGGAAGTGTAGTTGATTATTTATAGACACATGTAAATGTCTTATTGA 69641  
QY 1243 ----- 1243  
Db 69640 TTAAAAATTAATATCTTTTATTCTAGAGATTTTAGAAATTTGTACAATTAATAATAAAAA 69581  
QY 1243 ----- 1243  
Db 69580 TGCATTCCACGPTTAATCTGTATATCAAAAAAACCATTACTPAATATTACAGAAACATTAA 69521  
QY 1243 ----- 1243  
Db 69520 CTATTAGCTTCCTTTTGTACAAAAACTTAAAGTTATGATGTGTATATGAATGTGAT 69461  
QY 1243 ----- 1243  
Db 69460 ATGAAATATAGATACTATATATGTATATGTATATATAGTATGATATATGATGATGAATGTA 69401  
QY 1243 ----- 1243  
Db 69400 TAGTGTACTACTATATTTAGGATATCTGTATCCTACTTTTATATGTAAAAAGTATACTGTA 69341  
QY 1243 ----- 1243  
Db 69340 AATATATTCTAATTTCAATGAATAGCCTTCAAAATGGCATACAAATTTTATCATGTGAAC 69281  
QY 1243 ----- 1243  
Db 69280 ATAAATAACCATGTGATTGTTGGTATTAGGCTATTTCTCATATGTCAACTAGTATAAAT 69221  
QY 1243 ----- 1243  
Db 69220 AATGCCATTATAAGCATCCTTATTTCATTTTCGGTGAATCTCTGATTATATCTGAAGTTA 69161  
QY 1243 ----- 1243  
Db 69160 TATTTCTAGTCCCAAAATGAGTGGGTAAATGAGTGAACATTTTATGATCTTTTATATG 69101  
QY 1243 ----- 1243  
Db 69100 TACTAGTTAATGTCTCCAGCAGAAAGCCTATGTATTAGTTTTCATTTCTATCCCCAGTAT 69041  
QY 1243 ----- 1243  
Db 69040 ATGAGAGCATTTGATTTTATTTCACCTTTCCACCAATGTTACACACAGAGAAACATAT 68981

QY 1243 ----- 1243  
Db 68980 ACACAGACATACATACCTTCCCAAAATTAATAATTTACCTAAATCAACATAATTTGCAT 68921  
QY 1243 ----- 1243  
Db 68920 3GGTTTGGTTAGTAGTGATATGGACATATTTTATGAATATTTGTCTGTTTGTGTTCA 68861  
QY 1243 ----- 1243  
Db 68860 TTTTAAAGCATTTTCTCTCTTTTACTATTGTAATTTGATACAAAAATATATCATATCCC 68801  
QY 1243 ----- 1243  
Db 68800 CAAACAAAAATTTGAAAAAACCAGAAATCATGTGAAAAAACACAGATATCACCTATA 68741  
QY 1243 ----- 1243  
Db 68740 ATACCATGATCAGTTGATGATATATTAGCTGATATTACACTCTTGAGTATCATATGTGTG 68681  
QY 1243 ----- 1243  
Db 68680 TATATATTTTATATATATTCATATATTTTATTGCAAAATGTTATGCTTCCATATTTGTCG 68621  
QY 1243 ----- 1243  
Db 68620 TCCCTGCTCTTTTGATTAGCAATATTTTCTCAAGCTAAAGCATATATATTTTATACCAT 68561  
QY 1243 ----- 1243  
Db 68560 ATGATATACTATAATCTTATGCAACCAATCCCTGTCTATTGGCCATCTAGAAAAATCAAGT 68501  
QY 1243 ----- 1243  
Db 68500 TATTTTTCATTTTAAAGTAACATCAATTAATAATATCTTTAAATATATAGAAAAATATTTT 68441  
QY 1243 ----- 1243  
Db 68440 AAGCATATTTCTCAATATTTTCTCTCCAAAAATCTATTTTCTTATTGTATCAATATATTA 68381  
QY 1243 ----- 1243  
Db 68380 CCTGAAATTTTGGATTTCAACTATCTTAAGTTTAAATTTTATCAATATTTTATATATAAAA 68321  
QY 1243 ----- 1243  
Db 68320 TTTGCACTATATATGATATGAATGATACTATGCAGTAGATCTTATTTAGCGCTCTTTTAA 68261  
QY 1243 ----- 1243  
Db 68260 TATTTTGTGATATTAGCATTTGGAGATTTTCTCAAAATATAAAGTAGATGAATTCGCCAG 68201  
QY 1243 ----- 1243  
Db 68200 CTTTATGATCTTTAATAGTATATGAATAAAAATCATTTGTTTCAATATTTGAAAGAGA 68141  
QY 1243 ----- 1243  
Db 68140 ATAACTATAGCTATTTTGGGTAGTAAATTTGTTAAATCTTAAAAAGAAATGTCAGGACTAT 68081  
QY 1243 ----- 1243  
Db 68080 TGGCTTATTTCTCATTTTCTTGAGTCAATTTTGTGAAGGATGAAAGATGTATTAGTCAG 68021  
QY 1243 ----- 1243  
Db 68020 TTCTCATGCTGCTAATAAAGAAATACCAGACTGGGTAATTTTATAAAGGAAGAGGCTT 67961  
QY 1243 ----- 1243  
Db 67960 NATGAGCTACAGTTCCCATATGGCTTAATGAGGCCTCACAATCATGCTGAAGGCAAGGA 67901

QY 1243 ----- 1243  
Db 67900 GGAGCAAGCTCACATCTTACATGGTGGCAGCAAAAGAGCATGTGCAGGGAACTCCCAT 67841  
QY 1243 ----- 1243  
Db 67840 GTATAAACCATCAGATCTTGTGAGACTTATTCATCATCACAGAACACACAGGAAGA 67781  
QY 1243 ----- 1243  
Db 67780 CCCACCCTATGATTCATTAATCACTCCACAGGTCCCCACACACACACGTGGGAATTATG 67721  
QY 1243 ----- 1243  
Db 67720 GGAGCTACAGTTTGGATTTGGTGGGATACAGCCAAACCATAGCAAAAGAGTAGGATT 67661  
QY 1243 ----- 1243  
Db 67660 AAAACCATCTCTGAATTTATTTATATATAGCCATTCAATGATGCATCTGTATTGGT 67601  
QY 1243 ----- 1243  
Db 67600 AGGAGTAGAAGCTAGCTATAAAGTATTATTAAATGAGCAGGAGATTACAAAAATGACG 67541  
QY 1243 ----- 1243  
Db 67540 AGAAAAATAAACTCAAGATGTTTGGCTATGAGAGAGAGAAAGAGTTAAACACGGTAAC 67481  
QY 1243 ----- 1243  
Db 67480 GGAGAGGGCTTTATGACAAGGAGAAAGAAATTTTAAAGGGTGATGTGAAGTTA 67421  
QY 1243 ----- 1243  
Db 67420 TCAGGTTTGAGGACTTGGGATGACTTATCAGACCGGGGATGAGACGTTTAGATCAAAA 67361  
QY 1243 ----- 1243  
Db 67360 AAACCCTTTTGAATAGAAGCCAAATTTACCTAAAGTTAATTTATCTACTGAAACTGTAAC 67301  
QY 1243 ----- 1243  
Db 67300 TTTCTTACTATTACGGGAACTAAATCCATAGAGAAATCCCGAGATTCCTTTTGGT 67241  
QY 1243 ----- 1243  
Db 67240 CTTTATGTGGTGGTGGCAGAGGTAAGCAGAGGCTTAGAATATGAAAAACATTTAAC 67181  
QY 1243 ----- 1243  
Db 67180 AAAATAAAGAAATGAACATAAAAATGTTAAGATCAATCTTAAATATTGGAGATCCCTG 67121  
QY 1243 ----- 1243  
Db 67120 CAGACTTTTGGCAGCAAGTTTAAATGACCCCGGTATATATATCTGCAGTGAAC 67061  
QY 1243 ----- 1243  
Db 67060 AGTATTTCAGGAACCGCTTAAAGGAGGTAGAGATTGGGCTTATCCTGCATTTGTACAC 67001  
QY 1243 ----- 1243  
Db 67000 AGCCATGTGACGAGAGGTCTAGTCCAAAGCAATTTGTTCAATTTGGGAAAGAGAGTGT 66941  
QY 1243 ----- 1243  
Db 66940 GTAGTTATGATCTTAGTATACATCTTACAGGAGTAAAGTGAAGCAAGAAAGACTGA 66881  
QY 1243 ----- 1243  
Db 66880 TAATGGAGGACACAGAAGCAAGTGGTCTTATATAGATCAAACTTGAGCCAAATGAAC 66821  
QY 1243 ----- 1243

Db 66820 ATTAGTTATATTGGAGAAATTTGAAAAATGAGAGGGCAAGATAGGATGTTTGGGCTCAG 66761  
QY 1243 ----- 1243  
Db 66760 ATAGGACACTGAAATTTGAAAAATGAAATCAATAATGTTTGTGTTTCAACATCTAGAGTA 66701  
QY 1243 ----- 1243  
Db 66700 GAAATGAATATCTGGTTTGGCTCGCTGGAGATGTTTCAACCCCTTGAGTGTGACTAGGA 66641  
QY 1243 ----- 1243  
Db 66640 CAAACACAGTCATGATATAGGGTTTATTCACTTCACAGGTTCTTGTCTTAATCTATGTAA 66581  
QY 1243 ----- 1243  
Db 66580 CATTGAAATTCCTGCCTTCCCTTACCACCCATATTCATCAAGCACATTTGCTACTTTTCAA 66521  
QY 1243 ----- 1243  
Db 66520 GTTTTCTACTCAGCATTTCTCCCAACCTCTCTCAGGAGTAAAGACACCCATGTTTGG 66461  
QY 1244 ----- 1260  
Db 66460 TGCCCTTTGTAGATAGAAATAGAGTGGCTGGAGAGACACCGAGTGTCTCTGGGAGAGTGA 66401  
QY 1260 uiletrpHisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAl 1280  
Db 66400 GATCTGCACGAGGCTCTCGGCACAGTGTGTGATGACTCTCTGGACCTGCCCGGCG 66341  
QY 1280 aGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlase 1300  
Db 66340 GGAAGTGGTGTGTGACAGCTGGCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 66281  
QY 1300 rPheGlyGlnGlyThrGlyThrIletrpLeuaspAspMetAtqCysGlyAsnGluSe 1320  
Db 66280 GTTGGCCAGGGAACCTGGAACCATCTGGTGGATGATCGGTGCAAGGAATGAGTC 66221  
QY 1320 rPheLeutrPaspCysHisAlaLysProtrpGlyGlnSerAspCysGlyHisLysGluAs 1340  
Db 66220 ATTTCTATGGGACTGTACGCCAACCCCTGGGACAGAGTGTGTGGACACAGGAAGA 66161  
QY 1340 pAlaGlyValArgCys----- 1345  
Db 66160 TGCTGCGTGGGTGCTCTGGTAAAGTGTGAAAGGCTGCATTTAAACTGAGCCTCCCACT 66101  
QY 1345 ----- 1345  
Db 66100 AATGGTAGAACACTGGGAAACAGGCTCACTCAGAGAGTTGAGCTGAAGATAGGAT 66041  
QY 1345 ----- 1345  
Db 66040 TTCTTGTGAAGAATGAAAGCTATGTTAAACTAACGTTTTTCCACAACATAAATAAGATA 65981  
QY 1345 ----- 1345  
Db 65980 ATGTTCCACTGTTATCTAAATAGCCCCCATAATTTTGAGGGAAGAAAGGTGGATTGGC 65921  
QY 1345 ----- 1345  
Db 65920 CTTAGACTTTGAAGCAAGAGGATTCTGACTGATATTTTAAATGACATCCAGTATTACAG 65861  
QY 1345 ----- 1345  
Db 65860 GAGATGTTAAAGGATATTTAAATGACATCCGATAATCGCAGGAGATGTAAAAAGAAATA 65801  
QY 1346 ----- 1350  
Db 65800 CCAGTTCCTTTTGGTTCATTTACGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 65741  
QY 1350 euLysSerLeuAsnAlaSerSer----- 1357  
|||||

```
Db 65740 TGAATCACTGAATGCCTCCTCAGGTATAGTCTTCTGCTATACAGTGGCGACACAAATAC 65681
QY 1357 ----- 1357
Db 65680 CTAAGGCTACATGGAAGGTCAAGGGAGGCCAGAACTGCTATGATAATCAAGAAGACTG 65621
QY 1357 ----- 1357
Db 65620 GGGACTGCAGAAATATAGTCTGTGGTGGCGAGCCAGTGGTATCAATAGCTCTGTGACTG 65561
QY 1357 ----- 1357
Db 65560 CTTTGCTATGCAGTCCCTCTCTCTAGAAAAATCCCTTTTACATGCAGAGNAGGATAAGAC 65501
QY 1357 ----- 1357
Db 65500 TCTCATGTACCTCATAGATAATCTTCTAGATTCTCAAAATATCTTTATGAGACTGGC 65441
QY 1357 ----- 1357
Db 65440 AGAAACAGCTCAAGCTTCAGGAGAATCACTTAAACCCAGGAAGCAGAGATTGCAGTGAGC 65381
QY 1357 ----- 1357
Db 65380 CGAGATCGTCCAGTACACTCAGCCTGGGTGACAGAAATGAGACTCCATCTCAAAAAAAA 65321
QY 1357 ----- 1357
Db 65320 AAAAAAGAAAGAACAGCTCAGCTCAAGCCCTAGTATTTCTGATTCTGGTGTCTTAGT 65261
QY 1357 ----- 1357
Db 65260 AGTGAATGCAATGCAAGAGAAAGGTTAGTTATCAAAATCTCTGTATAAAATTTGAAGT 65201
QY 1357 ----- 1357
Db 65200 AATTTAGGCAATAGAAATATGGAAGATTTGGATTGCTTCCAGATAGTTTTCAT 65141
QY 1357 ----- 1357
Db 65140 CTAAATAGCTCTCAAGATTACGTACTGTTTCATCAATATGTTGAATGTATAGCCATTA 65081
QY 1357 ----- 1357
Db 65080 TTCTTCTCAGCAACTAGGAACACAGTCTTTTTCATAATGTATAAACCCACAAATTGATATC 65021
QY 1358 ----- GlyHisLeuAlaLeuIleLeus 1365
Db 65020 TACTAGTTAATATCAATGCTTAATTTCTATACAGGTCATTTAGCACTTATTTTAT 64961
QY 1365 erSerIlePheGlyLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArgV 1385
Db 64960 CCAGTATCTTTGGGCTCCTTCTCTGGTTCCTGTTATCTTATCTTCTCAGTGGTCCCGAG 64901
QY 1385 aLgInLySgInLyHisLeuProLeuArgVal----- 1395
Db 64900 TTCAGAAAACAAAACATCTGCCCTCCAGAGG-TGAGAGCCCACTAGGATTGTATAAAAT 64842
QY 1395 ----- 1395
Db 64841 GTGAGTTACCTGCCACCAAGGCCACACAGGATGGGGATCTTTGAGAAATGGAAGAAATG 64782
QY 1395 ----- 1395
Db 64781 CACTTGCCACTTTTGTAGTGAGATGGATTGGAAGTCCAGGCCCAATGCAGAAACATGTGG 64722
QY 1395 ----- 1395
Db 64721 AATGATCAAGAAATCAGCCAATCCACAATTTCCATTTACACAAATGTAATCTTTTAGAAGC 64662
QY 1395 ----- 1395
Db 64661 TGTAGCTTGCCCATACAGACGATGAGTCAGATAGGCAAAATATCTTTAGCTGATAGAAGGAG 64602
```

```
QY 1395 ----- 1395
Db 64601 GTAATTGTATGTTTGGTTCAGAACTCAGACTCCAGAAATAGCCATAGAAAATGTAGAA 64542
QY 1395 ----- 1395
Db 64541 GCATGCTGGAGGATTAAAGAACAAGAGTTTCTCTGGGTTTGAACACCACTTATATAACTTAC 64482
QY 1395 ----- 1395
Db 64481 TAGCTGTGTGGCCTTGGCTAGAACTACCTTTTCCAAACTCCAAATTTCCCTAATTTGTAAA 64422
QY 1395 ----- 1395
Db 64421 GATAATAATAGTCTTACCTCATCAGGTTAAGATAGTTAAATAAATTAATCATCTGTATGA 64362
QY 1395 ----- 1395
Db 64361 AAAATGGCCAGCACATAGTAAATACTATGTAAAGAACTTTGGTAAATAAAAGTGAAGAATA 64302
QY 1395 ----- 1395
Db 64301 TCAAGGGAGAAAAGAGGAGGATAAATTTTAAAAATGTAAAAAATTCACAGAGGGGTAGG 64242
QY 1395 ----- 1395
Db 64241 TGCTGAAAGAGCCATTTGGATTTTAGCTCAGAGGGATTAATATGTGTCAAACTTAGCAAT 64182
QY 1395 ----- 1395
Db 64181 AAGTGTCTTAATATAAGGGATAGGCCATTTGAATTTTGAATCCTTAACCATGAAGCATCCAT 64122
QY 1396 ----- SerThrArgArgArgGlySerLeuGluAsnLeuPheHisGluMetGluT 1413
Db 64121 TCTCAGTTTCAACCCAGAGAGGGGTTCTCTCGAGAGAAATTTATTTCCATGAGATGGAGA 64062
QY 1413 nrCysLeuLysArgGluAspProHisGlyThrArgThrSer 1426
Db 64061 CCTGCCTCAAGAGAGAGGAGGCCACATGGGACAAAGAACCTCA 64021
```

Search completed: May 12, 2003, 19:20:08  
Job time : 6250 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:24:52 ; Search time 52 Seconds  
(without alignments)  
5757.432 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDFGRCCHQON.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriopl:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8109	99.6	1453	4 Q9NR16	Q9nr16 homo sapien
2	3441.5	42.3	1151	4 Q07899	Q07899 homo sapien
3	3429	42.1	1116	4 Q07898	Q07898 homo sapien
4	3429	42.1	1156	4 Q07900	Q07900 homo sapien
5	3396.5	41.7	1149	4 Q07901	Q07901 homo sapien
6	3272	40.2	1121	11 Q99MX8	Q99mx8 mus musculus
7	2990	36.7	2403	4 Q9UGM2	Q9ugm2 homo sapien
8	2987.5	36.7	2426	4 Q9UGM3	Q9ugm3 homo sapien
9	2987	36.7	2412	4 Q9UJ57	Q9uj57 homo sapien
10	2982	36.6	2413	4 Q96DU4	Q96du4 homo sapien
11	2980	36.6	2413	4 Q9UKJ4	Q9ukj4 homo sapien
12	2448	30.1	1785	4 Q9Y211	Q9y211 homo sapien
13	2439.5	30.0	2153	5 Q97375	Q97375 strongyloce
14	2435	29.9	1785	4 Q9Y4V9	Q9y4v9 homo sapien
15	2246.5	27.6	2043	5 Q96943	Q96943 geodicia cydo
16	2156.5	26.5	1594	6 Q95218	Q95218 oryctolagus

17	1962.5	24.1	2083	11 Q60997	Q60997 mus musculus
18	1845	22.7	1957	11 Q9JMJ9	Q9jnj9 mus musculus
19	1594.5	19.6	1188	5 Q9NC87	Q9nc87 strongyloce
20	1586	19.5	710	6 Q95JES	Q95jes ovis aries
21	1431.5	17.6	749	5 Q9NC88	Q9nc88 strongyloce
22	1399	17.2	804	6 Q29113	Q29113 sus scrofa
23	1341	16.5	546	6 Q29111	Q29111 sus scrofa
24	1301	16.0	1075	5 Q9NC90	Q9nc90 strongyloce
25	1277	15.7	973	5 Q9NC89	Q9nc89 strongyloce
26	1219	15.0	369	6 Q29110	Q29110 sus scrofa
27	1205	14.8	420	6 Q95JD9	Q95jd9 ovis aries
28	1085.5	13.3	480	6 Q29112	Q29112 sus scrofa
29	1078.5	13.3	600	6 Q28910	Q28910 bos taurus
30	1048.5	12.9	1290	11 Q62827	Q62827 rattus norv
31	1041.5	12.8	470	13 Q98TD1	Q98td1 gallus gall
32	968.5	11.9	575	4 Q8WTU2	Q8wtu2 homo sapien
33	866.5	10.6	1036	5 Q97378	Q97378 strongyloce
34	763.5	9.4	437	6 Q29109	Q29109 sus scrofa
35	723.5	8.9	749	11 Q99JC8	Q99jc8 rattus norv
36	703	8.6	209	6 Q28881	Q28881 ovis aries
37	686.5	8.4	457	5 Q9NC91	Q9nc91 strongyloce
38	665	8.2	918	13 Q92098	Q92098 petromyzon
39	664.5	8.2	268	6 Q9TV20	Q9tv20 oryctolagus
40	664.5	8.2	277	6 Q9TV21	Q9tv21 oryctolagus
41	662.5	8.1	264	6 Q9TV22	Q9tv22 oryctolagus
42	650	8.0	505	4 Q96180	Q96180 homo sapien
43	632	7.8	754	11 Q91VN8	Q91vn8 mus musculus
44	620.5	7.6	665	11 Q91WN5	Q91wn5 mus musculus
45	603.5	7.4	592	4 Q8WWJ7	Q8wwj7 homo sapien

#### ALIGNMENTS

#### RESULT 1

Q9NR16 ID Q9NR16 PRELIMINARY; PRT; 1453 AA.  
AC Q9NR16;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Scavenger receptor cysteine-rich type I protein M160 precursor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20540099; PubMed=11086079;  
RA Gronlund J., Vitved L., Lausen M., Skjodt K., Holmskov U.;  
RT "Cloning of a Novel Scavenger Receptor Cysteine-Rich Type I  
Transmembrane Molecule (M160) Expressed by Human Macrophages.";  
RL J. Immunol. 165:6406-6415(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Gronlund J., Vitved L., Lausen M., Skjodt K., Holmskov U.;  
RL Submitt (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF264014; AAF91396.1; -  
DR InterPro: IPR001190; Srcr\_receptor.  
DR Pfam: PF00530; SRCR\_12.  
DR PRINTS; PR00258; SPERACTRCPTR.  
DR SMART; SM00202; SR; 11.  
DR PROSITE; PS00420; SRCR\_1; UNKNOWN\_5.  
DR PROSITE; PS0287; SRCR\_2; 12.  
KW Receptor; Signal.  
FT SIGNAL 1 42 POTENTIAL.  
SQ SEQUENCE 1453 AA; 159263 MW; 15445CE9868CF58A CRC64;

Query Match 99.6%; Score 8109; DB 4; Length 1453;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NMLPQNSWHIDFGRCCHQONLFSAVVTCILLNSCFLISSFNGTDLRLVNGDPCSGT 60

```
Db 1 MMLPQNSHIDFGRCCHONLFSAVVTCILLNSFLISSFNGTDLRLVNGDPCSGT 60
QY 61 VEVKFGQWGTCDGWNNTASTVVKQKQCPFSFAMRFGQAVTRHGKIWLDDVSCYGN 120
Db 61 VEVKFGQWGTCDGWNNTASTVVKQKQCPFSFAMRFGQAVTRHGKIWLDDVSCYGN 120
QY 121 ESALBECQREHSGHNCYHGEDVGNVCYGEANLGLRLVDGNNSCGRVEVKFQERWGTIC 180
Db 121 ESALBECQREHSGHNCYHGEDVGNVCYGEANLGLRLVDGNNSCGRVEVKFQERWGTIC 180
QY 181 DQGNLNTAAVVCROLGCPSSFISSGVNNSPAVLRIWLDDTLCOGNELALNCHRHGNG 240
Db 181 DQGNLNTAAVVCROLGCPSSFISSGVNNSPAVLRIWLDDTLCOGNELALNCHRHGNG 240
QY 241 NHDCSHNEVDVLTCTDSSDLRLVGGTNRCMGRVELKIQGRWGTVCCHKKNNAADVVC 300
Db 241 NHDCSHNEVDVLTCTDSSDLRLVGGTNRCMGRVELKIQGRWGTVCCHKKNNAADVVC 300
QY 301 KQLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFVWDCRHSGTVPNFDCLHQNDVSI 360
Db 301 KQLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFVWDCRHSGTVPNFDCLHQNDVSI 360
QY 361 CSDGADLELRADGNSNCGRVEVRIHQWWTICDQNKNEQALVVKQKQCPFSFVGSR 420
Db 361 CSDGADLELRADGNSNCGRVEVRIHQWWTICDQNKNEQALVVKQKQCPFSFVGSR 420
QY 421 RAKPSNEARDIWINISCTGNSALWDCITYDGKAKRTCFRRSDAGVICSADLRLVG 480
Db 421 RAKPSNEARDIWINISCTGNSALWDCITYDGKAKRTCFRRSDAGVICSADLRLVG 480
QY 481 AHSPCYGRLEVYKQEWGTVCHDRSTRNAAVVVKQKQCPHVFMTYFKEASGPIWL 540
Db 481 AHSPCYGRLEVYKQEWGTVCHDRSTRNAAVVVKQKQCPHVFMTYFKEASGPIWL 540
QY 541 DVVSCIGNESNIWDCHEHSGWGHKNCVHREDVITVCSGATWGLRVGNSNCGRLEVYF 600
Db 541 DVVSCIGNESNIWDCHEHSGWGHKNCVHREDVITVCSGATWGLRVGNSNCGRLEVYF 600
QY 601 QGRWGTVCDDGWNKAAAVVCSQDPCSSIIGMGLGNASTGYGKIWLDDVSCDGEDSLW 660
Db 601 QGRWGTVCDDGWNKAAAVVCSQDPCSSIIGMGLGNASTGYGKIWLDDVSCDGEDSLW 660
QY 661 SCRNNGWNNDCSHSEDEVGVCSDASDMLRVLGSSRCAGKVEVNVQAVGILCANGW 720
Db 661 SCRNNGWNNDCSHSEDEVGVCSDASDMLRVLGSSRCAGKVEVNVQAVGILCANGW 720
QY 721 MNIAEVVCRQLECGSAIRVSRPHEFTERTLHILMSNSGCTGGEASLWDCIRKWKOTACH 780
Db 721 MNIAEVVCRQLECGSAIRVSRPHEFTERTLHILMSNSGCTGGEASLWDCIRKWKOTACH 780
QY 781 LNWEASLICSARHQRPLVGADMPGSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNGD 840
Db 781 LNWEASLICSARHQRPLVGADMPGSGRVEVKHADTWRSVCDSDFSLHAANVLCRELNGD 840
QY 841 AILSVGDFHFGNGLTWAEKFCQSGSETHLALCPVHQHPEDTCHSREVGVVCSRYTDV 900
Db 841 AILSVGDFHFGNGLTWAEKFCQSGSETHLALCPVHQHPEDTCHSREVGVVCSRYTDV 900
QY 901 RLNVNKSQCDGVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGKYGIGERSV 960
Db 901 RLNVNKSQCDGVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGKYGIGERSV 960
QY 961 RVWGHFRHCLGNESLLDNCQMTVLGAPPCIRHNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
Db 961 RVWGHFRHCLGNESLLDNCQMTVLGAPPCIRHNTVSVICTGSLTQPLFPCLANVSDPYLS 1020
QY 1021 AVPEGSALICLEDKRLRLVGDGSRCAGRVEIYHDFWGTICDDGHDLSDAHVVCKLCCG 1080
Db 1021 AVPEGSALICLEDKRLRLVGDGSRCAGRVEIYHDFWGTICDDGHDLSDAHVVCKLCCG 1080
QY 1081 VAFNATVSAHFEGSGPIWLDDNLCTGFESHLLWQCPSGRWGQHDCRHKEDAGVICSEFTA 1140
```

```
Db 1081 VAFNATVSAHFEGSGPIWLDDNLCTGMESHLLWQCPSGRWGQHDCRHKEDAGVICSEFTA 1140
QY 1141 LRLYSETETESCAGLEVYFNGTWSVGRNRNITTAIAGIVCRQLCGGNGVVS LAPLSKT 1200
Db 1141 LRLYSETETESCAGLEVYFNGTWSVGRNRNITTAIAGIVCRQLCGGNGVVS LAPLSKT 1200
QY 1201 GSGFMVDDIOCPKTHISIWQCLSAWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260
Db 1201 GSGFMVDDIOCPKTHISIWQCLSAWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260
QY 1261 IWHAGSWGTVCCDDSDWDLAAEAEVVCQOLGCGSALALRDASFQCGTGTIWLDDMRCKGNES 1320
Db 1261 IWHAGSWGTVCCDDSDWDLAAEAEVVCQOLGCGSALALRDASFQCGTGTIWLDDMRCKGNES 1320
QY 1321 FLWDCCHAKPWGSDCGHKEDEAGVRCGSLKSLNASSGHLALILSIFGLLLVLFILFL 1380
Db 1321 FLWDCCHAKPWGSDCGHKEDEAGVRCGSLKSLNASSGHLALILSIFGLLLVLFILFL 1380
QY 1381 TWCRVQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDDTPNHGCCDASDT 1440
Db 1381 TWCRVQKQKHLPLRVSTRRRGSLLENLFHEMETCLKREDPHGTRTSDDTPNHGCCDASDT 1440
QY 1441 SLLGVLPASEATK 1453
Db 1441 SLLGVLPASEATK 1453

RESULT 2
Q07899
ID Q07899 PRELIMINARY; PRT: 1151 AA.
AC Q07899;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE M130 antigen cytoplasmic variant 1 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93380506; PubMed=8370408;
RA Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA Mason D.Y.;
RT "A new macrophage differentiation antigen which is a member of the
RT scavenger receptor superfamily."
RL Eur. J. Immunol. 23:2320-2325(1993).
DR EMBL: 222969; CAA80542.1; -
DR InterPro: IPR001190; Scrv_receptor.
DR Pfam: PF00530; SRCR; 9.
DR PRINTS: PR00258; SPERACTRCPTR.
DR SMART: SM00202; SR; 9.
DR PROSITE: PS00420; SRCR_1; 4.
DR PROSITE: PS50287; SRCR_2; 9.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 1151 AA; 124820 MW; 3CDBAD857B6B91 CRC64;

Query Match 42.3%; Score 3441.5; DB 4; Length 1151;
Best Local Similarity 55.2%; Pred. No. 4e-260;
Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

QY 364 GADLELRADGNSNCGRVEVRIHQWWTICDQNKNEQALVVKQKQCPFSFVGSRRAK 423
Db 41 GTDELRLVDGKNSCGRVEVKVQEWGTVCNNGWSMEAVSVICNQLCGPTAIRKAPGAN 100
QY 424 PSNEARDIWINISCTGNSALWDCITYDGKAKRT-CFRRSDAGVICSADLRLVGAH 482
Db 101 SSAGSGRIWMDHVSCKGNESALWDCIKHDGKHSNCTHQDAGVTCSDGNSLMRLTRGG 160
QY 483 SPCYGRLEVYKQEWGTVCHDRSTRNAAVVVKQKQCPHVFMTYFKEASGPIWLDD 542
```

Db 161 NMSGRIEIKFQGRWGTCDNFNDHASVTCROLEGSAVSFGSSNFGSGGPIWFD 220  
Qy 543 VSCIGNESNIWDCSHGKHNCHVREDVIVTCSDGTWGLRLVGGNRCGRLEVFQ 602  
Db 221 LICNGESALWNCCKHQGNCHDHAEDAGVICSKGADLSRLVDGYTECSGRLEVFQ 280  
Qy 603 RWGTVCDGWNKAAAVVCSOLDPCSSITIGMGLNASTGYGKIWLDDVSCDGDSDLSWC 662  
Db 281 EWGTCDDGWDSDYDAVACKQLGCPATAIGRNVASKGFGHIWLDVSCQGHPEAVQC 340  
Qy 663 RNSGWNDCSHSDVGVICSDASDMELRLVGGSSRCAGKEVNVQAGVILCANGCMN 722  
Db 341 KHEWGHYCNHNDAGVTCSDGDLRLRGGSRCAGTVEIQRLGKVCDRGWLK 400  
Qy 723 IAEVVCROLEGSAIRVSRPHEFTERTLHILMSNGTGGEGEASLWDCIRWEWKOTACHLN 782  
Db 401 EADVVCROLEGSAIKTSYVYSKIQTATNWLFLSSCNGNETSLWCKNQGGLTCDHY 460  
Qy 783 MEASLICASHPRLVGDADPCSGREVVEKADTWRSVCDSDFSLSHAANVLCRELNCDAI 842  
Db 461 EAKITCSAHPRLVGDGIPCSGRVEVKHGDWTGSIHSDSFLSLEAAVLCRELQCGTV 520  
Qy 843 SLVGDHFGKNGUTWAEKFCQEGSETHALCPITVQHPEDTCHSREVVGVCSRYTDVRL 902  
Db 521 SILGAHFGGNGQIAEAEFCQEGHESHLSCPAPRPGTCSHSRDVGVVCSRYTEIRL 580  
Qy 903 VNGKSQDGOVEINVLGHWSLCTHWDPEARVLCQLSCGTALSTTGGKYIGERSVRV 962  
Db 581 VNGKTPCEGRVELKLGAWSLCHSHWDIEDAHVLCQOLKCGVALSTPGGAREKNGQI 640  
Qy 963 WGRHFCGLNESLLDNCOMTVLGAHPCHTGNVTSVITGSLTOPFLPCLANVSDPYLSAV 1022  
Db 641 WRHMFHCTGTEQHMGDPCVPTALGASLPCSEOVASVICSGNQSLSSCSLGPTRPTI 700  
Qy 1023 PEGSALICLEDKRLRLVDGSRACGRVEIYHGGFWGTCDDGWDLSDAHVVVCQLGCGVA 1082  
Db 701 PEESAVACIESGQRLRLVNGGRCAGREIYHGGFWGTCDDGWDLSDAHVVVCQLGCGVA 760  
Qy 1083 FNATVSAHFEGSGPIWLDLNTCTGTESHLWQCPSRGQGHDRKEDAGVICSEFTALR 1142  
Db 761 INATGSAHFEGSGPIWLDLNTCTGTESHLWQCPSRGQGHDRKEDAGVICSEFTALR 820  
Qy 1143 LYSTETESACGRLEVPYNGTWGVSGRNNTTATAGIVCROLGCGGNGVWSLAPLSKTS 1202  
Db 821 LTSEASREACAGREVPYNGTWGVSGRNNTTATAGIVCROLGCGGNGVWSLAPLSKTS 880  
Qy 1203 GFMMVDIOCPKTHISTWQCLSPAWERRISSPAETWITCEDRIRVGGDTCCSGRVEIW 1262  
Db 881 IPMVDNVQCPKGPDTLWQCPSPSWERKLASPSSETWITCNKRLQEGPTSCSGRVEIW 940  
Qy 1263 HAGSWGTVCDSDWDLAEAVVVCQOLGCGSALALDRASFGGTTIWLDDMRCKGNESFL 1322  
Db 941 HGSWGTVCDSDWDLDAVVCQOLGCGPALKAFKEAFEGGTTIWLNEVCKGNESFL 1000  
Qy 1323 WDCHAKPWGSDCHGKEDAGVRCGSQLK-----SLNASSGHUALILSSIFGLLLVLF 1376  
Db 1001 WDCPARWGHSECHGKEDAAVNCNTDISVQKTPQKATTGRSSRSQSFIAVGLGLVLLAIF 1060  
Qy 1377 I--LFLTWCRVOKHPLRVTRRRGSLLENLPHMETCLKREDPHGTRTSDTTPHGC 1434  
Db 1061 VALEFLTKRRQROR---LAVSSRGENLVHQIYQIRREMNSCLNADDLMLNSEN--SHES 1115  
Qy 1435 EDASDTSLGV--LPASEATK 1453  
Db 1116 ADFSAAELISVSKFLPISGMEK 1137

RESULT 3  
ID Q07898 PRELIMINARY; PRT: 1116 AA.  
AC Q07898;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

01-MAR-2002 (Tremblrel. 20, Last annotation update)  
M130 antigen precursor.  
CD163.  
Homo sapiens (Human).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=93380506; PubMed=8370408;  
Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,  
Mason D.T.;  
"A new macrophage differentiation antigen which is a member of the  
scavenger receptor superfamily";  
Eur. J. Immunol. 23:2320-2325(1993).  
[2]  
SEQUENCE FROM N.A.  
Ritter M., Buehler C., Langmann T., Schmitz G.;  
"Genomic organization of the human CD163 gene";  
Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z22968; CAA80541.1; -  
DR EMBL; Y18388; CAB45233.1; -  
DR EMBL; Y18389; CAB45233.1; JOINED.  
DR EMBL; Y18390; CAB45233.1; JOINED.  
DR EMBL; Y18391; CAB45233.1; JOINED.  
DR EMBL; Y18392; CAB45233.1; JOINED.  
DR EMBL; Y18393; CAB45233.1; JOINED.  
DR EMBL; Y18394; CAB45233.1; JOINED.  
DR EMBL; Y18395; CAB45233.1; JOINED.  
DR EMBL; Y18396; CAB45233.1; JOINED.  
DR EMBL; Y18397; CAB45233.1; JOINED.  
DR EMBL; Y18398; CAB45233.1; JOINED.  
DR EMBL; Y18399; CAB45233.1; JOINED.  
DR EMBL; Y18400; CAB45233.1; JOINED.  
DR EMBL; Y18401; CAB45233.1; JOINED.  
DR EMBL; Y18402; CAB45233.1; JOINED.  
DR EMBL; Y18403; CAB45233.1; JOINED.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00530; Srcr; 9.  
DR PRINTS; PR00258; SPERACTRCPT.  
DR SMART; SM00202; SR; 9.  
DR PROSITE; PS00420; SRCR\_1; 4.  
DR PROSITE; PS00287; SRCR\_2; 9.  
KW Antigen; Signal.  
FT SIGNAL 1 40 POTENTIAL.  
FT CHAIN 41 1116 M130 ANTIGEN.  
SQ SEQUENCE 1116 AA; 120979 MW; E805BC80B52E4009 CRC64;  
Query Match 42.18; Score 3429; DB 4; Length 1116;  
Best Local Similarity 56.2%; Pred. No. 3.7e-259;  
Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;  
Qy 364 GADLELRADGNNCSGRVEVRIHEOMWTICDQWKNQALVVCQKLGCPFSVFSRRAK 423  
Db 41 GTDELRLVDGKNCGRVEVVKVQEEGTVCNNGMSAEVSVICNLQCPATAIGRNVAN 100  
Qy 424 PSNEARDIWNISICTGNESALWDCTYDGAART--CFRRSDAGVICSADLRLVGAH 482  
Db 101 SSAGSGLRWMDHVSRCGNESALWDCKHDGKHSNCTHQDAGVTCSDGSLNEMRLTRGG 160  
Qy 483 SPYGRLEVKYQEGWGTCHDRWSTRNAVVCQKLGCKPMHVFQMTYFKBASGPIWLD 542  
Db 161 NMSGRIEIKFQGRWGTCDNFNDHASVTCROLEGSAVSFGSSNFGSGGPIWFD 220  
Qy 543 VSCIGNESNIWDCSHGKHNCHVREDVIVTCSDGTWGLRLVGGSSRCGRLEVFQ 602  
Db 221 LICNGESALWNCCKHQGNCHDHAEDAGVICSKGADLSRLVDGYTECSGRLEVFQ 280  
Qy 603 RWGTVCDGWNKAAAVVCSOLDPCSSITIGMGLNASTGYGKIWLDDVSCDGDSDLSWC 662  
Db 281 EWGTCDDGWDSDYDAVACKQLGCPATAIGRNVASKGFGHIWLDVSCQGHPEAVQC 340  
Qy 663 RNSGWNDCSHSDVGVICSDASDMELRLVGGSSRCAGKEVNVQAGVILCANGCMN 722

```
Db 341 KHEWGKHYCHNEDAGVTCSDGSDLELRRLRGSGRCAGTVEVEIQRLRLGKVCODRGWLK 400
Qy 723 IAEVVCRLQEGSALRVSRPHFTERTLHILMSNGCTGGASLWDCIRWVKWQTAACHLN 782
Db 401 EADVVCRLQCGSALKTSYQVYSKIQATNTWLFSSCNGNETSLWDCNKGWGLTCDHY 460
Qy 783 MEASLICSARHQPRLVGDMPGSRVYKVDHATWRSVCDSDFSLHAANVLCRELNCDAI 842
Db 461 BEAKITCSAHREPRVLGGDIPCSGRVEYKHGDTWGSICDSDFSLEAASVLCRELQCGTW 520
Qy 843 SLVSGDHFEGKNGLTWAEKFCGSETHALALCPVQHEDTCIHSREVGVCSRYTDVRL 902
Db 521 SILGAHFEGBGNGQIWAEEFQCEGHESHLSCPAPRPEGTCSHSRDVGVCSTYETRL 580
Qy 903 VNGKSCQDQGVIEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTGGKYGIGERSVRV 962
Db 581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKNGQI 640
Qy 963 WGRHFLCNGNESLLDNCOMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLSAV 1022
Db 641 WRHMFHCTGTQEHMGDCPVTALGASLCPSEQVAVSICSGNSQSLSSCNSSSLGTRPTI 700
Qy 1023 PEGSALICLEDKRLRLVGDGRSACGRVEIYHDGFWGTTICDDGWLSDAHVVCOKLGCQVA 1082
Db 701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGWSGTTICDDSWLSDAHVVCRLGCGEA 760
Qy 1083 FNATVSAHFEGSGPILWDLNCTGESHLMQCPGSRGQHDCHKEKEDAGVICSFTALR 1142
Db 761 INATGSAHFEGGTGPIWLDENKCKESRIWQCHSHGWGQNCRHKEKEDAGVICSFTALR 820
Qy 1143 LYSETETESACRLEVFYNGTWGSGVRNITATAGIVCRLGCGENGVSAPLSKGTGS 1202
Db 821 LTSEASREACAGRLEVFYNGTWGSGVRNITATAGIVCRLGCGENGVSAPLSKGTGS 880
Qy 1203 GFWMVDDIQCPTKTHISIWQCLASAPWERRISSPAETWTTCEDIRVRGDDTECSRVEIW 1262
Db 881 IPMWVDDIQCPTKTHISIWQCLASAPWERRISSPAETWTTCEDIRVRGDDTECSRVEIW 940
Qy 1263 HAGSGTVCDDSDWDLAEAVVCCQQLGCGSALAAALDASFGQGTGTIMLDDMRCKGNESFL 1322
Db 941 HGGSGTVCDDSDWDLDDAAVVCQQLGCGPALKAFAEFGGTGPIWLVNCKGNESFL 1000
Qy 1323 WDCHAKPMGQSDCHKEKEDAGVICSQSLK-----SLNASSGHIALILUSSIFGLLLVLF 1376
Db 1001 WDCPARRGHSCEGHKEKEDAAVNCCTDISYQKTPQKATTCGRSRQSSFIAGVILGVVLLAIF 1060
Qy 1377 I-LFLTCRWQKOKHPLRVSTRRRGSLLENLPHMETCLKRED 1419
Db 1061 VALFELTKRRQRQR---LAVSSRGENVLHIOIYREMNCLNADD 1102
```

## RESULT 4

```
Q07900 ID Q07900 PRELIMINARY; PRT; 1156 AA.
AC Q07900;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE M130 antigen cytoplasmic variant 2 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=93380506; PubMed=8370408;
RA Law S.A., Micklem K.J., Shaw J.M., Zhang X.P., Dong Y., Willis A.C.,
RA Mason D.Y.;
RT "A new macrophage differentiation antigen which is a member of the
RL scavenger receptor superfamily."
RL Eur. J. Immunol. 23:2320-2325(1993).
DB EMBL; Z22970; CAA80543.1; .
```

```
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; SRCR; 9.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR; 9.
DR PROSITE; PS00420; SRCR_1; 4.
DR PROSITE; PS50287; SRCR_2; 9.
KW SIGNAL.
FT CHAIN 1 40 POTENTIAL.
FT CHAIN 41 1156 M130 ANTIGEN CYTOPLASMIC VARIANT 2.
SQ SEQUENCE 1156 AA; 125352 MW; 3BEEC22D805BD5C CRC64;
```

Query Match 42.1%; Score 3429; DB 4; Length 1156;  
Best Local Similarity 56.2%; Pred. No. 3.8e-259;  
Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

```
Qy 364 GADLELRADGNSNCSGRVEVRIHEQWWTICDNNKNQALVVCVKQLGCPFSVFCSSRAK 423
Db 41 GTDELRLVGENKSGRVEVYKVEEWTVCNNGMSMAVSVICNQLGCPAIPAGWAN 100
Qy 424 PSNEARDIWINISICTGNESALWDCTYDGKAKRT-CFRRSDAGVICSADKADLDRLVGAH 482
Db 101 SSAGSGRIWMDHVSCRGNESALWDCKHDGKGHSNCTHQQDAGVTCSDGSLNEMRLTRGG 160
Qy 483 SPCYGRLEVKYQEGMTVCHDRMSTRNAAVVCVKQLGCKPKMHVFMOTYFKEASGPIWLDD 542
Db 161 NMCSGRIEIKFQGRWGTVDNDFNIDHASVICRQLECGSAVSFGSSNFGEGSGPIWEDD 220
Qy 543 VSCIGNESINWDCSHGKHKHCHVREDIVVTCSDATWGLRLVCGSNCRSGRLEVYFQG 602
Db 221 LICNGESALNCKHQGWGKHNCHDAEDAGVICSAGDLSURLVDVGTCTECGRLEVRFG 280
Qy 603 RWGTVCDDGWSKAAAVVCSQDLCPSSTIGMLGNASTGYGKIWLDDVSCDGEDSLWSC 662
Db 281 EMGTICDDGWSYDAAVACKQLGCPATAIGRVNASKGFGHILWDSVSCOGHEPAVWQC 340
Qy 663 RNSGWNDCSHSDEVDVGCSDASDMELRLVGGSSRCAGKVEVNVQGVAGVILCANGWGN 722
Db 341 KHEWGKHYCHNEDAGVTCSDGSDLELRRLRGSGRCAGTVEVEIQRLRLGKVCODRGWLK 400
Qy 723 IAEVVCRLQEGSALRVSRPHFTERTLHILMSNGCTGGASLWDCIRWVKWQTAACHLN 782
Db 401 EADVVCRLQCGSALKTSYQVYSKIQATNTWLFSSCNGNETSLWDCNKGWGLTCDHY 460
Qy 783 MEASLICSARHQPRLVGDMPGSRVYKVDHATWRSVCDSDFSLHAANVLCRELNCDAI 842
Db 461 BEAKITCSAHREPRVLGGDIPCSGRVEYKHGDTWGSICDSDFSLEAASVLCRELQCGTW 520
Qy 843 SLVSGDHFEGKNGLTWAEKFCGSETHALALCPVQHEDTCIHSREVGVCSRYTDVRL 902
Db 521 SILGAHFEGBGNGQIWAEEFQCEGHESHLSCPAPRPEGTCSHSRDVGVCSTYETRL 580
Qy 903 VNGKSCQDQGVIEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTGGKYGIGERSVRV 962
Db 581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKNGQI 640
Qy 963 WGRHFLCNGNESLLDNCOMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLSAV 1022
Db 641 WRHMFHCTGTQEHMGDCPVTALGASLCPSEQVAVSICSGNSQSLSSCNSSSLGTRPTI 700
Qy 1023 PEGSALICLEDKRLRLVGDGRSACGRVEIYHDGFWGTTICDDGWLSDAHVVCOKLGCQVA 1082
Db 701 PEESAVACIESGQLRLVNGGGRCAGRVEIYHEGWSGTTICDDSWLSDAHVVCRLGCGEA 760
Qy 1083 FNATVSAHFEGSGPILWDLNCTGESHLMQCPGSRGQHDCHKEKEDAGVICSFTALR 1142
Db 761 INATGSAHFEGGTGPIWLDENKCKESRIWQCHSHGWGQNCRHKEKEDAGVICSFTALR 820
Qy 1143 LYSETETESACRLEVFYNGTWGSGVRNITATAGIVCRLGCGENGVSAPLSKGTGS 1202
Db 821 LTSEASREACAGRLEVFYNGTWGSGVRNITATAGIVCRLGCGENGVSAPLSKGTGS 880
Qy 1203 GFWMVDDIQCPTKTHISIWQCLASAPWERRISSPAETWTTCEDIRVRGDDTECSRVEIW 1262
```



"Molecular cloning and characterization of the mouse CD163 homologue, a highly glucocorticoid-inducible member of the scavenger receptor cysteine-rich family.";  
 RL Immunogenetics 53:170-177(2001).  
 DR EMBL: AF274883; AAK16065.1; -  
 DR MGD: MGI:2135946; Cdl63.  
 DR InterPro; IPR001190; Srcr\_receptor.  
 DR Pfam; PF00530; SRCR; 9.  
 DR PRINTS; PR00258; SPERACTRCPTR.  
 DR SMART; SM00202; SR; 9.  
 DR PROSITE; PS00420; SRCR\_1; UNKNOWN\_3.  
 DR PROSITE; PS50287; SRCR\_2; 9.  
 KW Receptor; Signal.  
 FT SIGNAL 1 38 POTENTIAL.  
 SQ SEQUENCE 1121 AA; 120889 MW; 04428DF18AD26187 CRC64;

Query Match 40.2%; Score 3272; DB 11; Length 1121;  
 Best Local Similarity 52.4%; Pred. No. 6.9e-247;  
 Matches 582; Conservative 174; Mismatches 332; Indels 22; Gaps 9;  
 QY 320 GSDVYVLDGVSQGNESFLWDCRHSHTVNFDCIHLONDVSVICSDGADL--ELRLADGNN 377  
 DB 3 GHRMYLLGAGSGPGCKREV---HLGFFVAVSLLSASAVTNAPGEMKKELRLAGGENN 58  
 QY 378 CSGRVEVRTHQWWTICDNNKNEQALVYVCKLQCPFFSVFSGRRRAKPSNEARDIWNIS 437  
 DB 59 CSGRVELKHDKWGTVCNSGMSNNEVSVVCOQLGCPSTIKALGWANSAGSGYIMDKVS 118  
 QY 438 CTGNESALWDCYTDGKAKRTCFRRSDAGVICSADLRLV-GAHSYCYGRLEVKYQGE 496  
 DB 119 CTGNESALWDCKHDKGKHNCHEKADAGVTCSDGNSLEMLRVNSAGHCLGRVELKFGOK 178  
 QY 497 WCTVCHDRWTRNAAVCKQLGCKPMVFMTYFKASGPIWLDVSCIGNESINWDCPE 556  
 DB 179 WGTVCDDNFSDKHASVICKQLGCGSAISFGSAKLGAGSGPIWLDLACNGESALWDC 238  
 QY 557 HSGWKHNCVHREDVIVTCSDATWGLRLVGSNRCRLEVFYQGRWGTVCDDGWSKA 616  
 DB 239 HSGWKHNCVHREDVIVTCSDATWGLRLVGSNRCRLEVFYQGRWGTVCDDGWSKA 616  
 QY 617 AAVVCSQDLPSSIIIMGLNASTGYKILWLDVSCDGSDESLMSRNSGWNDCSHSE 676  
 DB 299 ASVVCKQLGCPATAISGRVNASGSGQIWLNDNISCEGHEATLWCKHQQKHYCHHRE 358  
 QY 677 DVGVTCSDASDELRLVSGRSCAGKVEVNVGAGILCANGCMNTAEVVCRLQEGSA 736  
 DB 359 DAGVTCSDGADLEURLVGGSGRAGIVEVEIQKLTGKMCRCRGWTLADADVVCRLQEGSA 418  
 QY 737 IRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWEKQTAHLNMEASLICSAROPR 796  
 DB 419 LQTAQKIYKGTATNLTWLPSCNGNETTFWQCKNWQWGLSCDNFEAKVTCSGHREPR 478  
 QY 797 LVGADMPGSGREVYKHADTWRSVCDSPSLHAANVLCRELNCGBAISLSVGDHFGKNG 856  
 DB 479 LVGGEIPGSGREVYKHADTWRSVCDSPSLHAANVLCRELNCGBAISLSVGDHFGKNG 856  
 QY 857 TWAEKFOCGSETHALCPIVQHPEDTCHHSRVEGVVCSRYTDVRLVNGKSCDQGVN 916  
 DB 539 TWAEKFOCGSETHALCPIVQHPEDTCHHSRVEGVVCSRYTDVRLVNGKSCDQGVN 916  
 QY 917 VLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGGYIGERSVRVWGHFRHCLGNESLL 976  
 DB 599 TLGAWGLPCLSHHWDHAHLVLCQLKCGVAQSIPEGAHFGKAGQVWVSHFCHTGTTEHI 658  
 QY 977 DNCQMTVLGAPCTHNGTVSVICTGSLTQPLFC--LANVSDPYLSAVPEGSALICLEDK 1034  
 DB 659 GDCLMTALGAPCTSEGQVAVSICSGNSQTLPCSSLSPVQTT--SSVIPSEVPCTASG 717  
 QY 1035 RLRLVGDGSRCAGRVEIYHDFGWTICDDGWDLSDAHVCQKLGCGVAFNATVSAHREG 1094  
 DB 718 QLRLVGGGRCAGREVVECHESGWTVCDDNDWMTDANVVCQLDCCGVAINATGAYFEG 777  
 QY 1095 SGPIWLDLNLCTGTESHLWQCPSPRGWGHDCRHKEDAGVICSEPTALRLXSETETES 1154

DB 778 AGAIWLDEVICTGKESHVQCHSHGWRHCRHKEDAGVICSEFWSRLTNEAHKENC 837  
 QY 1155 RLEVFYNGTWSVGRNRNITTAAGIVCRLQGGGNGVYSLAPLSKTSGSGFMVWDIOCPK 1214  
 DB 838 RLEVFYNGTWSVGRNRNITTAAGIVCRLQGGGNGVYSLAPLSKTSGSGFMVWDIOCPK 897  
 QY 1215 THISWQCLSPAPWERRISSP-AEETWITCEDRIRVRGDDTECSGRVEIWHAGSWGTVCDD 1273  
 DB 898 GVDILWQCPSPFWKORQASPSQSWIICDNKIRLOEGHTDCSGRVEIWHAGSWGTVCDD 957  
 QY 1274 SWDLAEVVCQQLGCGSALALRDASFGQGTGTTWLDMDRCCKNESFLWDCPAPWQS 1333  
 DB 958 SWDLNDKAVVCKLQCGQAVKALKEAAGPGTGPILWNEIKRCNGNESSLWDCPAPWQS 1017  
 QY 1334 DCGHKEDAGVRCGSLKSLNASSGHLALILSSIFGLLLVLFILFTWC-----RVQKQK 1389  
 DB 1018 DCGHKEDASIOCLPKMTSESHGHTGHTPTALLVCGAILLVLLVFLWTLKRRRIOR-- 1075  
 QY 1390 HPLRVSTRRRGRSLEENLFHEMETCKRKED 1419  
 DB 1076 --LTVSSRGEVLIHQVQYQEMDS--KADD 1100  
 RESULT 7  
 QYUGM2 PRELIMINARY; PRT: 2403 AA.  
 AC QYUGM2;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE DMBT1/8kb.1 protein.  
 GN DMBT1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG;  
 RX MEDLINE=20065089; PubMed=10597221;  
 RA Mollenhauer J., Holmskov U., Wiemann S., Krebs I., Herberitz S.,  
 RA Madsen J., Kioschis P., Coy J.F., Poustka A.;  
 RT "The genomic structure of the DMBT1 gene: evidence for a region with  
 RT susceptibility to genomic instability.";  
 RT Oncogene 18:6233-6240(1999).  
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
 DR EMBL: AJ243224; CAB63942.1;  
 DR InterPro; IPR000859; CUB\_domain.  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR InterPro; IPR001190; Srcr\_receptor.  
 DR Pfam; PF00431; CUB; 2.  
 DR Pfam; PF00530; SRCR; 14.  
 DR Pfam; PF00100; zona\_pellucida; 1.  
 DR PRINTS; PR00258; SPERACTRCPTR.  
 DR SMART; SM00042; CUB; 2.  
 DR SMART; SM00202; SR; 14.  
 DR SMART; SM00241; ZP; 1.  
 DR PROSITE; PS01180; CUB; 2.  
 DR PROSITE; PS00420; SRCR\_1; UNKNOWN\_13.  
 DR PROSITE; PS50287; SRCR\_2; 14.  
 DR PROSITE; PS00682; ZP\_DOMAIN; UNKNOWN\_1.  
 FT CHAIN 26 2403 DMBT1/8KB.1 PROTEIN.  
 SQ SEQUENCE 2403 AA; 259573 MW; 9026222DE7399AD38 CRC64;

Query Match 36.7%; Score 2990; DB 4; Length 2403;  
 Best Local Similarity 40.5%; Pred. No. 2e-224;  
 Matches 619; Conservative 177; Mismatches 504; Indels 230; Gaps 23;

QY 44 TDLERLVNGDGPSCGVTEVKFGQGWGTVCDDGWNITASTVCKQLGCPFSFAM---FRF 100  
 DB 230 SSLALRLVNGDGRGRVREVLRYGSGWGTVCDDYWDITNDANVVCRLGCGWAMSAPGNAQF 289



Qy	937	LCROLSCGTALSTGKGYIGERSVRVWGHRRHFCIGNESLLDNCOMTVLGAPCPIHGNTVS	936
Dy	1172	VCROLGCGWAMSPAGNARFGQSGPITVLDDVRCSCGHESYLWSCPHNGWLSHNCGHEDAG	1231
Qy	997	VICTGSLTQPLFPCLANVSDPYLSAVPEGSAIICLEDKRLRLVDGDSRCAGRVEIYHDGF	1056
Dy	1332	VICASQSQPT-PSPTWTPTSHASTAGSESSL-----ALRLVNGDRCQGRVEVLYRGS	1284
Qy	1057	WGITICDDGMDLSAHVVCOKLGGVAFNAFSAHFGEGSGPIWLDDNLCTGTESHLMQCP	1116
Dy	1285	WGTVCDDYDWTNDANVVCROLGCGWATSAPGNARFGQSGPILVDDVRCSCGHESYLWSCP	1344
Qy	1117	SRGWHQDCHKEDAGVICS-----EFTALRLYSETETESC	1152
Dy	1345	HNGWLSHNCGHEDAGVICSASQSQPT-PSPTWTPTSHASTAGSESSLALRLVN--GGDRC	1402
Qy	1153	AGRLVEFYNGTWSGVRNRNITTAIGIYVCROLGCGENGWVSLAPLSK---TGSGFMMVDD	1209
Dy	1403	QGRVEVLYRSGWGTVCDDYDWTNDANVVCROLGCG---CWATSAPGNARFGQSGPILVDD	1459
Qy	1210	IQCPKTHISTWQCLSPW-----ERRISSPAETWIT-----CEDR	1245
Dy	1460	VRCSCGHESYLWSCPHNGWLSHNCGHEDAGVICSASQSQPT-PSPTWTPTSHASTAGSEST	1519
Qy	1246	IRVR---GGDTESGRVETWHAGSGTVCDDSDWDLAEAEVVCQOLGCGSALAAALRDASF	1302
Dy	1520	LALRLVNGD-RCRGRVEVLYQSGWGTVCDDYDWTNDANVVCROLGCGWAMSPAGNAQFG	1578
Qy	1303	QGTGTIWLDDMRCKGNESFLWPCHPKQSGDCGHKEDAGVRCSGQSILKS	1352
Dy	1579	QGSQPILVDDVRCSCGHESYLWSCPHNGWLSHNCGHEDAGVICSAAQSQS	1628

RESULT 9

Q9UJ57

ID

Q9UJ57

PRELIMINARY;

PRT; 2412 AA.

AC

Q9UJ57;

DT

01-MAY-2000 (TREMBLrel. 13, Created)

DT

01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT

01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE

DMBT1/8kb.2 protein precursor.

GN

DMBT1.

OS

Homo sapiens (Human).

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CX

NCBI\_TaxID=9606;

RN

[1]

RP

SEQUENCE FROM N.A.

RC

TISSUE=TRACHEA.

RX

MEDLINE=99415938; PubMed=10485905;

RA

Holmskov U., Mollenhauer J., Madsen J., Vitved L., Groenlund J.,

RA

Tornøe I., Kilem A., Reid K.B., Poustka A., Skjodt K.;

RT

"Cloning of gp-340, a putative opsonin receptor for lung surfactant

RT

protein D.";

RL

Proc. Natl. Acad. Sci. U.S.A. 96:10794-10799(1999).

CC

-1- SIMILARITY: CONTAINS 2 CUB DOMAINS.

DR

EMBL; AJ243212; CAB56155.1; -.

DR

HSSP; P29392; 1SFP.

DR

InterPro; IPR000859; CUB\_domain.

DR

InterPro; IPR001507; Endoglin/CD105.

DR

InterPro; IPR001190; Srcr\_receptor.

DR

Pfam; PF00431; CUB; 2.

DR

Pfam; PF00530; SRCR; 14.

DR

Pfam; PF00100; zona\_pellucida; 1.

DR

PRINTS; PR00258; SPERACTRCPTR.

DR

SMART; SM00042; CUB; 2.

DR

SMART; SM00202; SR; 14.

DR

SMART; SM00241; ZP; 1.

DR

PROSITE; PS01180; CUB; 2.

DR

PROSITE; PS00420; SRCR\_1; UNKNOWN\_13.

DR

PROSITE; PS0287; SRCR\_2; 14.

DR

PROSITE; PS00682; ZP\_DOMAIN; UNKNOWN\_1.

KW

Signal.

```

Db 1187 DDVRCSGHESYLWSCPHNGWLSHNCGHHEHDAGVICSASQSPPTPSPDWTPTSHASTAGSE 12146
Qy 900 - - - - - VRLVNGSKSCDGGVEINVLGHWSLCDTHWDPEDARVLCRLSCGTALTSTGGKYI 955
Db 1247 SSLALRVNGGDRQCGRVEVLYRGSWGTVCCDDYWDTNDANVVCRLGCGWATSAPGNARF 13066
Qy 956 GERSVRVWGRHFHCLNGESLLDNCQMTVLGAPPCIHNGTVSVICTGSLTQPLPCLANVS 10151
Db 1307 GQSGGPVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEHDAGVICSASQSPPT - PSPDTPW 13655
Qy 1016 DPYLSAVPEGSSALICLEDKRLRLVDCGRCACGRVEIYHDGFGNGTTCDDGWLDSLDAHVVCQ 10753
Db 1366 TSHASTAGSESSL - - - - - ALRVNGGDRQCGRVEVLYRGSWGTVCCDDYWDTNDANVVCYR 14191
Qy 1076 KIGGVAFNATVSAHFGEGSGPTLWDLNCTGTESHLMOCPSRGGHGDCHRHKEHDAGVIC 11353
Db 1420 QLGGCWATSAPGNARFGQSGPVLDDVRCSGHESYLWSCPHNGWLSHNCGHHEHDAGVIC 14799
Qy 1136 SEF - - - - - TALRLYSETEDESCARLEVFYNGTWGSGVRGN 11717
Db 1480 SAFQSQPPTPSPDWTPTSHASTAGSESTLALRLVN - - - - - GGDRCRGRVEVLYQSGMGTVCCDY 15373
Qy 1172 ITTAIACIVCRLCGENGWVSLAPLUSK - - - - - TGSQPMWVDDTCQPKTHISIHQCLSAWP - 12273
Db 1538 WDTNDANVVCRLGCG - - - - - GWAMSPAGNAQFGQSGPVLDDVRCSGHESYLWSCPHNGWL 15949
Qy 1228 - - - - - ERRITSSPAETWIT - - - - - CEDRIWR - - - - - GGDTECSGRV 12559
Db 1595 SHNCGHHEHDAGVICSASQSQSPPTPDTWTNLNLPALTVGSESSLALRVNGGD - RCRGRV 16533
Qy 1260 EIWHAGSWGTVCCDSDWDLAEAEVVCQGLGSAALRALDRASFQGTGTITLWDDMRCKGNE 13191
Db 1654 EVLYRGSWGTVCCDSDWTNDANVVCRLGCGWAMSPAGNARFGQSGPVLGDRVCSGNE 17133
Qy 1320 SFLWDCHAKPWGSDCGHCKEDAGVRCGSGSLKSLNASSGH 1359
Db 1714 SYLWSCPHKGWLTHNCGHHEHDAGVICSATQINSTTTTDMWH 1753

RESULT 10
Q96D04 PRELIMINARY; PRT; 2413 AA.
ID AC Q96D04
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DMBT1/8kb.2 protein precursor.
GN DBMT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Mollenhauer J.;
RT "Major subforms of DMBT1 are gastrointestinal mucins that display
RL extensive alternative splicing and differential protein targeting.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297935; CAC4122.1; -
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 14.
DR Pfam; PF00100; zona_pellucida; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS00420; SRCR_1; UNKNOWN_13.
DR PROSITE; PS50287; SRCR_2; 14.
DR PROSITE; PS00682; ZP_DOMAIN; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.

```

FT	CHAIN	26	2413	DMBT1/8KB.2 PROTEIN.
SQ	SEQUENCE	2413 AA;	260600 MW;	99A449C5F4F60728 CRC64;
	Query Match	36.6%;	Score 2982;	DB 4; Length 2413;
	Best Local Similarity	40.1%;	Pred. No. 8.7e-24;	
	Matches 618;	Conservative 176;	Mismatches 506;	Indels 240; Gaps
QY	44	TDLELRVNGDGCPCSTVEVKFGQGWGTVCDDGWNNTASTVVCVKQLGCGFFSFAM---	FRF 100	
DB	230	SSLALRVNGDRCRGRVEVLYRSGWTVCCDDYWDTDNDANVVCRLQCGWAMSPGNAQF	289	
QY	101	GQAVTRHGKIWDDVGYCNGESALWECQHRGSHNCHYHGEDVGYCYG-----	149	
DB	290	GQG---SGPIVLDVRCSGHESYLMSCPNHGLTINCCHSEDAVGICSAPOSRPTSPDPT	346	
QY	150	-----EANLGLRLVDGNNCSGRVEYKFOERWGTICDDGWNLENTAAVVCRLQGC	198	
DB	347	WPTSHASTAGPESSALRLVNGDRCQGRVEVLYRSGWTVCCDDSWDTSANVVCRLQGC	406	
QY	199	PSFISGGVNVSPAVLRPIWDDILCOGNELALWNCRHRCGNCHDCSHNEDVLTICY---	255	
DB	407	GWATSPGNARFQGGSGPIVLDVRCSGYESYLMSCPNGHNSCHQHSEDAVGICSAH	466	
QY	256	-----DSSLRLVGGTNRCKMRVELKTOGRWGTVCCHHKWNA	294	
DB	467	SRSTPSPDTLPTITLPASTVGSSESLALRVNGDRCQGRVEVLYRSGWTVCCDSWDTN	526	
QY	295	AADVCKQLCGCGTALHAGLPHLOSCSDVWLDGYVSCSGNESFLWDCRHSGTVNFDCLHQ	354	
DB	527	DANVVCRLQCGWAMUAPGNARFGQGGSPIVLDVRCSGNESYLMSCPNGHNSCHNC	586	
QY	355	NDVSVICS-DGADLELRLDAGNNCSGRVEYRIHQWWTICDQNMKNQALVWCQLCGCP	413	
DB	587	EDAGVICSPSESSALRLVNGDRCQGRVEVLYRSGWTVCCDSWDTNDANVVCRLQGC	646	
QY	414	FSVFGSRRAKPSNEARDIWINISCTGNESALWDCTYDGKAKRTCFRRSDAGVICS---	469	
DB	647	WATSPGNARFQGGSPIVLDVRCSGHESYLMSCPNGHNSCHNCGHEDAGVICSAAQS	706	
QY	470	-----DKADLDRLVGAHSPCYGRLEVKYQGEWGTVCCHDRWSTRN	509	
DB	707	RSTPGPDTLSTITLPSTVGSSESLRLVNGSDRCQGRVEVLYRSGWTVCCDSWDTND	766	
QY	510	AAVVCQLGCGPMHVGMYTFKESAGPIWLDVSCIGNESINWDCSHSGKHCNCHVRE	569	
DB	767	ANVVCRLQCGWATSPAGNARFGQGGSPIVLDVRCSGHESYLMSCPNGHNSCHNC	826	
QY	570	DVLTVICS-----GDATW-----GLRLVGGSCSCGRLEYVYFQGRGTV	607	
DB	827	DAGVICSQSQRPTSPDTWPTWPTSHASTAGSESSALRLVNGDRCQGRVEVLYRSGTV	886	
QY	608	CDDGWNKAAAVVCQLDPCPSIIIGMGLGNASTGYGKIWLDVDCDGDSEDLWSCN	667	
DB	887	CDDSWDTSANVVCRLQCGWATSPAGNARFGQGGSPIVLDVRCSGYESYLMSCP	946	
QY	668	GNDCSHSEDEVGIVCSDA-----SMELRLVGGSSRCAGKV	703	
DB	947	LSHNCQHSDEDAVICSAAHWSWTPSPDTLPTITLPASTVGSSESLALRVNGDRCQGRV	1006	
QY	704	EYNVQGAIVLCANQWGMNIAVVCRLQECGSALRVSRPHFTERTLHILMNSNGCTGGE	763	
DB	1007	EVLYQSGWTVCCDSWDTNDANVVCRLQCGWAMSPAGNARFGQGGSPIVLDVRC	1066	
QY	764	ASLWDCIRPWEKQTACHLMEANSLICSAGR-OP-----RLVG	799	
DB	1067	SYLWSCPNGHNSCHSEDAVGICSAQSQRPTSPDTWPTWPTSHASTAGSESSALRLVN	1126	
QY	800	ADWPCSGRVEVKHADTVRSVCDSDFSLHAANVLCRELNCGDAISLSVGDHFGKGLTWA	859	
DB	1127	GGDRCQGRVEVLYRSGWTVCCDDYWDTDNDANVVCRLQCGWAMSPAGNARFGQGG	1186	
QY	860	EXFOCEGSETHALCPVIOHPEDTCTIHSREVGVCYSRYTD-----	899	

[illegible]

Qy	860	EKFCGEGSETHLALCPVQHPEDTCIHREVVGVVCSRYTD-----	899
Db	1187	DDVRCGCHIESYLWSCPHNGWLSHNCGHHEDAGVICSASQSQPTSPDTPWTSHTAGSE 1246	
Qy	900	----VRLVNGKSCQDQVEINVGLHWGSLCFDHPEDARVLCQLSCGTALSTTGKXI 955	
Db	1247	SSALRLVNGDRCQRCQREVLVYRGSWGTVCDYDWDNDANVVCYQLGCGWATSAPGNARF 1306	
Qy	956	GERSVRWGHRHFCILGNESLLDNCONQTVLGAPPCIHGNTVTSICGTSUTQPLFFCLANVS 1015	
Db	1307	GGCGSGPIVLDVRCSCGHSYLSWCPHNGWLSHNCGHHEDAGVICSASQSQPT-PSDTPW 1365	
Qy	1016	DPYLSAVPEGSGALICLEDKRLRLVDGDSRCAGRVEIYHDSFGWGTICDDGWLSDAHVVCQ 1075	
Db	1366	TSHASTAGSESSL-----ALRLVNGDRCQRCQREVLVYRGSWGTVCDYDWDNDANVVC 1419	
Qy	1076	KLGGVAFNATVSAHFGEGSGPIVLDLNLCTCTESHLMQCPSRGWGQHDCHRHKEDAGVIC 1135	
Db	1420	QLGCGWATSAPGNAREGGSGPIVLDVRCSCGHSYLSWCPHNGWLSHNCGHHEDAGVIC 1479	
Qy	1136	-----SEFT-ALRLYSTETESCAGRLEVYNGTWSGVGRN- 1171	
Db	1480	SASQSQPTSPDTPWTSRSTAGSESTLALRLVN--GGDCRGREVLVYQCSWGTCVDDY 1537	
Qy	1172	ITTAIGATVCRQLCGGSENGVWSLAPLUSK---TGSFGFMVDDIQCPKTHISTWQCLSPW- 1227	
Db	1538	WDTNDANVVCRLGCG---GWAMSPAGNAQFGQSGPIVLDVRCSCGHSYLSWCPHNGWL 1594	
Qy	1228	-----ERRISSPAETWLT-----CEDRIYR--CGDTECSGRV 1259	
Db	1595	SHNCGHHEDAGVICSAAQSOSTPRPDTWLTWLPALTVGSSSSIALRLVNGGD-RCKGRV 1653	
Qy	1260	EINHGSGWGTVCDDSDWDLAEVVCQOLGCGSALALRDASFGQGTGTIWLDDMRCKGNE 1319	
Db	1654	EVLVYRGSWGTVCDSDWDTNDANVVCRLGCGHMSAPGNAREGGSGPIVLDVRCSCGNE 1713	
Qy	1320	SFLWDCHAKPWGSDCGCHKEDAGVRCSGOSKSLNASSGH 1359	
Db	1714	SYLWSCPHKGWLTHNCGHHEDAGVICSATQINSTTTDWMH 1753	
RESULT- 12			
ID	Q9Y211	PRELIMINARY; PRT; 1785 AA.	
AC	Q9Y211;		
DT	01-NOV-1999 (TrEMBLrel. 12, Created)		
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)		
DE	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	DMBT1.		
GN	DMBT1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RX	SEQUENCE FROM N.A..		
RA	MEDLINE=20017478; PubMed=10551316;		
RA	Takeshita H., Sato M., Shiwa H.O., Sema S., Sakurada A., Hoshi M.,		
RA	Hayashi Y., Tagawa Y., Ayabe H., Horii A.;		
RT	"Expression of the DMBT1 gene is frequently suppressed in human lung		
RT	cancer."		
RL	Jpn. J. Cancer Res. 90:903-908(1999).		
CC	-1- SIMILARITY: CONTAINS 2 CUB DOMAINS.		
DR	EMBL; AB020851; BAA78577.1;		
DR	EMBL; AB020812; BAA78577.1; JOINED.		
DR	EMBL; AB020813; BAA78577.1; JOINED.		
DR	EMBL; AB020814; BAA78577.1; JOINED.		
DR	EMBL; AB020815; BAA78577.1; JOINED.		
DR	EMBL; AB020816; BAA78577.1; JOINED.		
DR	EMBL; AB020817; BAA78577.1; JOINED.		
DR	EMBL; AB020818; BAA78577.1; JOINED.		
DR	EMBL; AB020819; BAA78577.1; JOINED.		
DR	EMBL; AB020820; BAA78577.1; JOINED.		

DR EMBL; AB020821; BAA78577.1; JOINED.  
DR EMBL; AB020822; BAA78577.1; JOINED.  
DR EMBL; AB020823; BAA78577.1; JOINED.  
DR EMBL; AB020824; BAA78577.1; JOINED.  
DR EMBL; AB020825; BAA78577.1; JOINED.  
DR EMBL; AB020826; BAA78577.1; JOINED.  
DR EMBL; AB020827; BAA78577.1; JOINED.  
DR EMBL; AB020828; BAA78577.1; JOINED.  
DR EMBL; AB020829; BAA78577.1; JOINED.  
DR EMBL; AB020830; BAA78577.1; JOINED.  
DR EMBL; AB020831; BAA78577.1; JOINED.  
DR EMBL; AB020832; BAA78577.1; JOINED.  
DR EMBL; AB020833; BAA78577.1; JOINED.  
DR EMBL; AB020834; BAA78577.1; JOINED.  
DR EMBL; AB020835; BAA78577.1; JOINED.  
DR EMBL; AB020836; BAA78577.1; JOINED.  
DR EMBL; AB020837; BAA78577.1; JOINED.  
DR EMBL; AB020838; BAA78577.1; JOINED.  
DR EMBL; AB020839; BAA78577.1; JOINED.  
DR EMBL; AB020840; BAA78577.1; JOINED.  
DR EMBL; AB020841; BAA78577.1; JOINED.  
DR EMBL; AB020842; BAA78577.1; JOINED.  
DR EMBL; AB020843; BAA78577.1; JOINED.  
DR EMBL; AB020844; BAA78577.1; JOINED.  
DR EMBL; AB020845; BAA78577.1; JOINED.  
DR EMBL; AB020846; BAA78577.1; JOINED.  
DR EMBL; AB020847; BAA78577.1; JOINED.  
DR EMBL; AB020848; BAA78577.1; JOINED.  
DR EMBL; AB020849; BAA78577.1; JOINED.  
DR EMBL; AB020850; BAA78577.1; JOINED.  
DR InterPro; IPR000859; CUB\_domain.  
DR InterPro; IPR001507; Endoglin/CD105.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00431; CUB; 2.  
DR Pfam; PF00530; SRCR; 9.  
DR Pfam; PF00100; zona\_pellucida; 1.  
DR PRINTS; PR00258; SPERACTRCPTR.  
DR SMART; SM00042; CUB; 2.  
DR SMART; SM00241; ZP; 1.  
DR PROSITE; PS01180; CUB; 2.  
DR PROSITE; PS00420; SRCR\_1; UNKNOWN\_8.  
DR PROSITE; PS0287; SRCR\_2; 9.  
DR PROSITE; PS00682; ZP\_DOMAIN; UNKNOWN\_1.  
SQ SEQUENCE 1785 AA; 193991 MW; 38B2363F95226EB0 CRC64;

Query Match  
Best Local Similarity 30.1%; Score 2448; DB 4; Length 1785;  
Matches 534; Conservative 156; Mismatches 413; Indels 374; Gaps 28;

QY 43 GTD--LELRVNGDPCSGTVEVFCQGWGTVCDDGNTTASTVVCQKQCGPPSFAM--- 97  
DB 95 GDSGLALRLVNGDRCGRVEILYRGSGTVCDDSDWTDNDANVVCRLQCGWAMSAPGN 154

QY 98 FRFGQAVTRHGKIWLDDVSCYGNESALWECOHREWGSHNCYHGEDVGNYC----- 148  
DB 155 AWFQGG---SFIALLDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVCSAAPOQSTLR 211

QY 149 -----GBANLGLRLVNGNSCGRVEVKFQERWGTICDDGWNINLTAAYV 192  
DB 212 PESWPVRISPPVPTTEGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDNDANV 271

QY 193 CRLQCGSPSSFTSSGVNNSPAVLRIWLDDILCQNELALNCRHGNHDCSHNEVDTL 252  
DB 272 CRLQCGWAMSAPGNAOFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAV 331

QY 253 TCY-----DSSDLRLVNGTNRCMGRVCLKIOGRWGTVCCHK 290  
DB 332 ICAPQSRPTSPDPTWPTSHASTAGSESSALRLVNGDRCGRVEILYRGSGTVCDD 391

QY 291 WNNAAADVVCQKCGGTALHFAPLHQLQSGSDVVLWDGVSCSGNESFLMPCRHISGTYNFD 350  
DB 392 WDTSDANVVCRLQCGWAMSAPGNAOFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSH 451

QY 351 CLHONDVSVICSDGADLELRADGNSNCSGRVEVRIHQWWTICQNKWNEQALVVCQK 410  
DB 452 CQHSEDAGVCSAA-----HSWST----- 470

QY 411 GCPFSVFGSRRAPSNARDIWIINSISCTGNESALWDCITYDGKAKRKTFRRSRDAGVICSD 470  
DB 471 -----PSPDTLPITLTPASTVGSSE----- 491

QY 471 KADLDRLVLGAHSPCYGRLEVYQGEWGTVCHDRKSTNAAVVCQKQCGKPMHVFQMTY 530  
DB 492 ---LALRLVNGDRCGRVEILYRGSGTVCDDSDWTDNDANVVCRLQCGWAMSAPGNAR 548

QY 531 FKEASGPIWLDDVSCIGNESNIWDCHEGSHGKHCNVHREDVLTVCSDGA-----TW- 581  
DB 549 FGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAVICSASQSRPTSPDPTW 608

QY 582 -----GLRLVGSNRCSGRLEVYFQGRWGTVCDDGWNKSKAAAAYVCQKQDPCS 628  
DB 609 TSHASTAGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDNDANVVCRLQCGW 668

QY 629 SLIIGMLGNASTGYGKIWLDDVSCDGEDSLWSCNRSWGNNDCHSHSDGVICS----- 683  
DB 669 AMSAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAVICSASQ 728

QY 684 -----DADMELRLVGGSSRCAGKVEVNVQGVAVGILCANGWGNIAEV 726  
DB 729 PTPSPDTWPTSHASTAGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDNDANV 788

QY 727 VCRLEGSAIRVREPHFTERTLHILMSNGCTGGASLWDCIRWEWKOTACHLNMEAS 786  
DB 789 VCRQCGWATSAPCNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAV 848

QY 787 LICSAHR-QPRLVGADMPCSRVEVKHADTWRSVCDSDSFLHAANVLCRELNCGDAISLS 845  
DB 849 VICSASQSP-----TPSPDTWPT----- 867

QY 846 VGDHFGKNGLTWAEKFCGESETHALCPVQHPEDTCIHSREVGVVCSRYTQVRLVNG 905  
DB 868 -----SRASTAGSESTAL-----RLVNG 886

QY 906 KSCQDQGOVEINVLCHWGLCDTHWDPEDARVLCRLQSCGTALSTTGGKYIGERSVRVWGH 965  
DB 887 GDRCRGRVEILYQSGWGTVCDDYWDNDANVVCRLQCGWAMSAPGNAOFGQSGPIVL 946

QY 966 RFHCLGNESLIDNCOMTVLGAPCIHGNTVSVICTGSLTO--PLFPCLANVSDPYLSA 1023  
DB 947 DVRCSGHESYLWSCPHNGWLSHNCGHSEDAVICSASQSTPRPDTWLTNLPALT 1006

QY 1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWLSDAHVVCQKLGCGVAF 1083  
DB 1007 ESSL-----ALRLVNGDRCGRVEILYRGSGTVCDDSDWTDNDANVVCRLQCGW 1059

QY 1084 NATVSAHFQSGPIVLDDILNCTCTESHMQCPSRGQCHDCRKHEDAGVICSFEFTALRL 1143  
DB 1060 SAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHSEDAVICS---ATQI 1116

QY 1144 YSET-----ETESCAGRLEVY-NGTWGS----- 1166  
DB 1117 NSTTDDWHPTTTTARPSNCSGFL--FYASGTFSSPSYPAYYPNNAKCVWEIEVNSGY 1174

QY 1167 ---VGRNI-----TTAAGIVC---ROLGCGENVVSL----AP 1196  
DB 1175 RINLGFNLKLEAHNCSFDYVEIFDGLNSLLKRICNDTRQIFTSSYNRMTIHRSD 1234

QY 1197 LSKTGSFM-WVDIQCPKTHISIWQCLSPAWERRISSPAEETWITCEDRIRVGGDTEC 1255  
DB 1235 ISQNTGFLAWYN-----SFPSDATL-----RLVNLNSYGLC 1267

QY 1256 SGRVEIHWAGSWGTVCCDSDWDLAEAVVCQKQCGSALALRDASFCQGTCTIWLDDMRC 1315  
DB 1268 AGRVEIYHGWTGTVCCDSDWTLQEAEEVVCRLQCGRAVSALGNAYFGSGSGPITLDDVEC 1327



01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DMBT1/6kb.1 protein precursor.  
GN DMPT1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
RN RP SEQUENCE FROM N.A.  
RC TISSUE=LUNG;  
RX MEDLINE=97434209; PubMed=9288095;  
RA Mollenhauer J., Wiemann S., Scheuren W., Korn B., Hayashi Y.,  
RA Wilgenbus K.K., von Deimling A., Poustka A.;  
RT "DMBT1, a new member of the SRCR superfamily on chromosome 10q25.3-  
RT q26.1 is deleted in malignant brain tumours.";  
RL Nat. Genet. 17:32-39(1997).  
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
DR EMBL; AJ000342; CAA04019.1; -;  
DR InterPro; IPR000859; CUB\_Domain.  
DR InterPro; IPR001507; Endoglin/CD105.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00431; CUB; 2.  
DR Pfam; PF00530; SRCR; 9.  
DR Pfam; PF00100; zona\_pellucida; 1.  
DR PRINTS; PR00258; SPRACTRCPTR.  
DR SMART; SM00042; CUB; 2.  
DR SMART; SM00241; SR; 9.  
DR SMART; SM00241; ZP; 1.  
DR PROSITE; PS01180; CUB; 2.  
DR PROSITE; PS00420; SRCR\_1; UNKNOWN 8.  
DR PROSITE; PS0287; SRCR\_2; 9.  
DR PROSITE; PS00682; ZP\_DOMAIN; UNKNOWN\_1.  
DR SIGNAL.  
KW SIGNAL.  
FT CHAIN 1 25 POTENTIAL.  
FT CHAIN 26 1785 DMBT1/6KB.1 PROTEIN.  
SQ SEQUENCE 1785 AA; 193941 MW; 7B1F8D47EA82092 CRC64;  
  
Query Match 29.94; Score 2435; DB 4; Length 1785;  
Best Local Similarity 36.19; Pred. No. 3.3e-181;  
Matches 533; Conservative 155; Mismatches 415; Indels 374; Gaps 28;  
  
Qy 43 GTD---LERLVNGDPCSGTVEVFGQGMGTVCDDGWNTTASTVVKCOLGCPESFAM--- 97  
Db 95 GDSGLALRLVNGDRCGRVEILFRSGWGTVCDDSWDTNDANVVCRLGCGWMSAPGN 154  
Qy 98 FRFGQAVTRHGKIWLDDSYCSGNESALWECQHRWEGSHNCHYHGEDVGYNCY----- 148  
Db 155 AWFQGG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVCSAAQPSFLR 211  
Qy 149 -----GEANLGLRLVDGNNSCGRVEYKFOERWGTICDDGWNLTAAV 192  
Db 212 PESVPVRIIPPVPTGESSLALRLVNGDRCGRVEILFRSGWGTVCDDYWDNDANV 271  
Qy 193 CRLGCPSSFTSSGVNPAVLRLPWLDDILCOGNELALWNCRHRGWNHDCSHNEDVTL 252  
Db 272 CRLGCGWMSAPGNAQFGQSGPIVLDVDRCSGHESYLWSCPHNGWLTNCHGSEDAGV 331  
Qy 253 TCY-----DSSDLRLVGGTNRGMRVVELKIQGRWGTVCCHK 290  
Db 332 ICSPQRSPTSPDPTWPTSHASTAGPESLALRLVNGDRCGRVEILFRSGWGTVCDD 391  
Qy 291 WNNAAADVVCOLGCGTALHFAGLPHLOGSDVWLDGVCSGNESHFLWDCRHSCTVNF 350  
Db 392 WDTSDANVVCRLGCGWATSAPGNARFGQSGPIVLDVDRCSGVESYLWSCPHNGWLSH 451  
Qy 351 CLHNDVSVICSDGADLELRADGNSNCGRVEIRHQQWTTICDQWKNQALVVCQKL 410  
Db 452 COHSEDAGVCSAA-----HSWT----- 470  
Qy 411 GCPFSVGRRAKPSNEARDIWNISICTGNESALWDCYDYGKARTCFRSDAGVICS 470  
Db 471 -----PSPDTLPTITLPAVSGESS----- 491

Qy 471 KADLDLRLVGAHSPCYGRLEVYQGEWGTVCCHDRWSTNRNAAAVCKQLGCGKPMHVFGMTY 530  
Db 492 ---LALRLVNGDRCGRVEILFRSGWGTVCDDSWDTNDANVVCRLGCGWMSAPGNAR 548  
Qy 531 FKEASGPIWLDVSCIGNESINWCEHSGKHNCHVREDVIVTCSGDA-----TW- 581  
Db 549 FGQSGPIVLDVDRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSASQSRTSPDTPW 608  
Qy 582 -----GLRLVGGSNRCSGRLEVYFOGRWGTVCDDGWNSKAAAAYVCSOLDCPS 628  
Db 609 TSHASTAGSESLALRLVNGDRCGRVEILFRSGWGTVCDDYWDNDANVVCRLGCGW 668  
Qy 629 SIIMGWLGNASTGYKIWLDDVSCDGBSDLSWSCRNSGWNNDCSHSDEGVICS----- 683  
Db 669 AMSAPGNARFGQSGPIVLDVDRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSASQ 728  
Qy 684 -----DASDMELRLVGGSSRCACKVEVNVQGAIVILCANGWGNIAEV 726  
Db 729 PTPSPDPTWPTSHASTAGSESLALRLVNGDRCGRVEILFRSGWGTVCDDYWDNDANV 788  
Qy 727 VCRQLEGSAIRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWENKQATACHLNMEAS 786  
Db 789 VCRQLGCGWATSAPGNARFGQSGPIVLDVDRCSGHESYLWSCPHNGWLSHNCGHEDAG 848  
Qy 787 LICSAHR-QPRLVGADMPGSRVEVVKHADTWRSVCDSDFSLHAANVLCRELNCDAISLS 845  
Db 849 VICASASQSP-----TPSPDPTWPT----- 867  
Qy 846 VGDHFGKNGLTWAEKFOEGSETHALCPVQHPEDTCHHSREVGVCSRYTVDRLVNG 905  
Db 868 -----SRASTAGSESLAL-----RLVNG 886  
Qy 906 KSQCDGQVEINVLGHWSGLCTHMDPEDARVLCRLSCGTALSTGGYKIGERSVRYWGH 965  
Db 887 GDRGRGRVEILYQSGWGTVCDDYWDNDANVVCRLGCGWMSAPGNAQFGQSGPIVLD 946  
Qy 966 RPHCLGNESLLDNQMTVLGAPPCIHGNTVSVICTGSLTQ--PLFPCLANVSDPILSAPV 1023  
Db 947 DVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSAAQSQTTPRDTWLTNLPALTVGS 1006  
Qy 1024 EGSALICLEDKRLRLVDGSDRCAGREIYHDGFWGTICDDGWLDSDAHVVCQKLGCGVAF 1083  
Db 1007 ESSL-----ALRLVNGDRCGRVEILFRSGWGTVCDDSWDTNDANVVCRLGCGW 1059  
Qy 1084 NATVSAHFGESGPIWLDLNCGTGTEHLMOCPSRGWQHDCRHKEDAGVICSSETALRL 1143  
Db 1060 SAPGNARFGQSGPIVLDVDRCSGNESYLWSCPHNGWLSHNCGHEDAGVICS---ATQI 1116  
Qy 1144 YSET-----ETESCAGRLEVY-NGTWGS----- 1166  
Db 1117 NSTTDDWHPHTTTTARPSNCGGLF--FYASGTFSPSPYPAYPNNKAKVWEIEVNSGY 1174  
Qy 1167 ---VGRRI-----TTAGIVC---ROLGCGENGVSIL---AP 1196  
Db 1175 RINGFSLNLEAHNCSFDYVEIFDGSINSLLLGKICNDTRQIFTSSYNTMTHFRSD 1234  
Qy 1197 LSKTGSFPM-WVDIIQCPKTHISIWQCLASAPWERRISSPAEETWTCEDRIVRGDTEC 1255  
Db 1235 ISFQNTGFLAWN-----SPPSDATL-----RLVNLNLSYGL 1267  
Qy 1256 SGRVEIWAGSWGTVCDDSWDLAEAVVCOQLGCGSALAALRDASFGOGTGIWLDLMRC 1315  
Db 1268 AGREIYHGGWGTVCDDSWTIQAEAVVVCRLGCGRAVSALGNAYFGSGSPITLDDVEEC 1327  
Qy 1316 KGNESFLWDCWQSDCGHCKEDAGVRCSGQSLS 1352  
Db 1328 SGTESTLWQCRNRGWFSHCNHREDAGVICSGNHLST 1364

RESULT 15  
O96943  
ID O96943 PRELIMINARY; PRT; 2043 AA.



Job time : 71 secs

---



transmembrane molecule (M160), expressed by human macrophages.  
7. Immunol. 165 (11), 6406-6415 (2000)

JOURNAL  
MEDLINE  
PUBMED  
11086079  
2 (bases 1 to 4579)  
Gronlund, J., Vitved, L., Lausen, M., Skjodt, K. and Holmskov, U.  
Direct Submission  
Submitted (05-MAY-2000) Department of Immunology & Microbiology,  
Institute of Medical Biology, University of Southern Denmark,  
Odense University, Winsloe Parken 21.1, Odense 5000, Denmark

Location/Qualifiers  
1. .4579  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q13.3"  
1. .25  
26. .4387  
/codon\_start=1  
/product="scavenger receptor cysteine-rich type 1 protein  
M160 precursor"  
/protein\_id="AAF91396.1"  
/db\_xref="GI:9652087"  
/translation="MWLPONSWHIDPGRCCCHQNLFSAYVTCILLNSCFLISSFNFT  
DLELRVNGDPCSTVEYKFGQMGTCVDDGWTNTASTVVCQKQCPSPFAMFRFG  
AYTRKGIWLDVSCYGNESALWEGCHREMGSHNCVHGDEGVNCGYGEANLGLRLVIG  
NNSCSRGVEYKFOERWGTICDDGNWLNATAVVCROLGCPSPFISGVNPAVLRLVIG  
LDLILCGNELALNCHRGCHGHDICSHNEDVLTLCYDSDLELRGLVGTNMGKRGVE  
LKIQRWGTVCYCHHKNNAAADVVCQLGCGCTALHAGLPHLQSGSDVWLDGVSCSGN  
ESFLDGRHSIGVFNEDLQNDVSVICSDGADLELRADGNSNCSGRVRETHEQWMT  
ICDQNNKNEALVVCQKQCPSPFSGRRAKESNEARDIWNISICTGNESALWQYIT  
DGAARTGCFRRSDAGVICSADKADLRLVGAHSPGYGRLEVYQGEWGVCHDRWSTR  
NNAVKQILGCGCKPMHVFGMTYFKKASGPILWDDVSCIGNESNIWDCSHSGKHNVC  
HREPVITCSGATWGLRLVGGSNRCSGRLEVYFOGRWGTVCDDGWNKAAAVVCOSL  
DCEPSIILGMLGNASTGYCKIWLDDVDCDGEDSLWSCNCSGWNDCSHSEDEVGVIC  
SDASDMELRLVGGSNRCSGRLEVYVQGVAGVILCANGMNIAYVVCQKQCPSPFAMFRFG  
REPRTKRLHLMSNCSGTGGEASLWDCIRWEKQATACLHNAESALICSAHRQPLV  
GADMPCSGRVEYKHADTWKVSQSDFSLHNAVNLRELNCDAISLVGHDFGKGNGL  
TWAERFCQEGSETHALCP IVQHPEDTCHSREVGVVCSYRSDTVRLVNGKSCDQGOVE  
INVLGHWSLCTDTHDPEDARVQLQCSGTALSTGGKYICGRSVRVNMGHFRHCLGN  
ESLNDLCONMTVLGAPPCITGNTVSVICTSLTPQLPLCLANSDPVLSAVPFGSALIC  
LEDKRLRLVDGSRCSGRVEIYHDGFMTGICDGDWLDSDAHVVCOKLGGCVAFNATVS  
AHFGSGSGILWLDLNCCTMESHLQAQCSRGWGHDRKHEDAGVLCSEPTALRLYSE  
TETSCSAGLEVYFYNGTSGVGRNRNITTAIGVLCQCGENGVSVLAPLSKTGSGF  
MWVDDICPKTHSIWQLCSAPWERRISITAEETITCEDRIVRVGGDTECSGRVEIM  
HAGSWGTCHQDSDWDLAEAEVVCQKQCGSALAARDASFGQGTGTIWLDDMRCKNGNES  
FLWDGHAKPWGSDCGHEDAGVRCSGOSKINASGRRLALILSSIFGLLLPVLFI  
FLTWCVQKQKHLPLRVSTVRRRRSGLEENLFHEMETCLRKREDPHGTRTSDDTNPHGDC  
ASDTSILGLVPASEATK"  
26. .151  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 1"  
470. .790  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 2"  
791. .1111  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 3"  
1112. .1432  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 4"  
1433. .1753  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 5"  
1754. .2074  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 6"  
2075. .2395  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 7"  
2396. .2710  
/note="Region: scavenger receptor cysteine-rich domain  
SRCR 8"

sig\_peptide  
misc\_feature  
152. .469  
misc\_feature  
470. .790  
misc\_feature  
791. .1111  
misc\_feature  
1112. .1432  
misc\_feature  
1433. .1753  
misc\_feature  
1754. .2074  
misc\_feature  
2075. .2395  
misc\_feature  
2396. .2710

Db 619 TCCATCTCTTTATTTCTCTGGAGTTGTTAATAGCCCTGCTGTATTGCGCCCATTTG 678  
Qy 681 GCTGATCACATTTTATCCAGGGAATGAGTTGCACTCTGGAAATTCAGACATCTG 740  
Db 679 GCTGATGACATTTTATCCAGGGAATGAGTTGCACTCTGGAAATTCAGACATCTG 738  
Qy 741 ATGGGGAATCATGACTGCAGTCACAATGAGGATGTCACATTAACCTTGTATGATAG 800  
Db 739 ATGGGGAATCATGACTGCAGTCACAATGAGGATGTCACATTAACCTTGTATGATAG 798  
Qy 801 TGATCTTGAATAGCTTTGATGAGGAACTAACCGCTGTATGGGGAAGTAGAGCTGAA 860  
Db 799 TGATCTTGAATAGCTTTGATGAGGAACTAACCGCTGTATGGGGAAGTAGAGCTGAA 858  
Qy 861 AATCAAGGAAGTTGGGGAGCCGTATGCCACCATAAAGTGGAACTGCTGAGCTGATGT 920  
Db 859 AATCAAGGAAGTTGGGGAGCCGTATGCCACCATAAAGTGGAACTGCTGAGCTGATGT 918  
Qy 921 CGTATGCAAGCAGTTGGGATGGAACCGCACTTCACCTGCTGCTGCTCAATTTGCA 980  
Db 919 CGTATGCAAGCAGTTGGGATGGAACCGCACTTCACCTGCTGCTGCTCAATTTGCA 978  
Qy 981 GTCAGGCTCTGATGTTGATGGCTGATGGTGTCTCTGCTCCCGTAATGATCTTTCT 1040  
Db 979 GTCAGGCTCTGATGTTGATGGCTGATGGTGTCTCTGCTCCCGTAATGATCTTTCT 1038  
Qy 1041 TTGGACTGACAGACATTCGGAACCGTCAATTTTGGACTGCTCTCATCAAAACGATGTGTC 1100  
Db 1039 TTGGACTGACAGACATTCGGAACCGTCAATTTTGGACTGCTCTCATCAAAACGATGTGTC 1098  
Qy 1101 TGTGATCTGCTCAGATGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAAACAATG 1160  
Db 1099 TGTGATCTGCTCAGATGAGCAGATTTGGAACTGCGACTAGCAGATGGAAGTAAACAATG 1158  
Qy 1161 TTCAGGAGTAGAGGTGAGAAATCATGAACTGCTGGAGTATGTCAGCAACATG 1220  
Db 1159 TTCAGGAGTAGAGGTGAGAAATCATGAACTGCTGGAGTATGTCAGCAACATG 1218  
Qy 1221 GAAAGTAAGCAAGCCCTTGTGGTTGTAAGCAGTACGATGCTCCGTTGAGCTTTGG 1280  
Db 1219 GAAAGTAAGCAAGCCCTTGTGGTTGTAAGCAGTACGATGCTCCGTTGAGCTTTGG 1278  
Qy 1281 CAGTCTGCTGCTAACTAGTAATGAACTAGACATTTGGATAAAGCAGATATCTTG 1340  
Db 1279 CAGTCTGCTGCTAACTAGTAATGAACTAGACATTTGGATAAAGCAGATATCTTG 1338  
Qy 1341 CACTGGGAATGATGACTGCTGAGTATGAGTACATATGATGGAAGCAAGCAACATG 1400  
Db 1339 CACTGGGAATGATGACTGCTGAGTATGAGTACATATGATGGAAGCAAGCAACATG 1398  
Qy 1401 CTTCCGAAGATCAGATGCTGAGTAAATTTGTTCTGATAAGGAGATCTGGACCTAAGCT 1460  
Db 1399 CTTCCGAAGATCAGATGCTGAGTAAATTTGTTCTGATAAGGAGATCTGGACCTAAGCT 1458  
Qy 1461 TGTGGGCTCATAGCCCTGTTATGGAGATGAGGTGAAATACCAAGAGAGTGGGG 1520  
Db 1459 TGTGGGCTCATAGCCCTGTTATGGAGATGAGGTGAAATACCAAGAGAGTGGGG 1518  
Qy 1521 GACTGTGTGATGACATGAGCAGACAGGAAGTACAGCTGTTGTGTGAACAATTTGG 1580  
Db 1519 GACTGTGTGATGACATGAGCAGACAGGAAGTACAGCTGTTGTGTGAACAATTTGG 1578  
Qy 1581 ATGTGGAAGCTATGCAATGTTGTTGATGACCTATTTTAAAGAGCATCAGGACCTAT 1640  
Db 1579 ATGTGGAAGCTATGCAATGTTGTTGATGACCTATTTTAAAGAGCATCAGGACCTAT 1638  
Qy 1641 TTGGCTGATGACGTTCTTGCATTTGGAATCAGTCAAAATCTGGGACTGTGGAACAG 1700  
Db 1639 TTGGCTGATGACGTTCTTGCATTTGGAATCAGTCAAAATCTGGGACTGTGGAACAG 1698  
Qy 1701 TGGATGGGAAGCAATATTTGTACACAGAGAGATGATGTAAGCTGCTCAGGTGA 1760  
Db 1699 TGGATGGGAAGCAATATTTGTACACAGAGAGATGATGTAAGCTGCTCAGGTGA 1758

Qy 1761 TGCAACATGGGCTGAGGCTGTTGGGGGAGCAACCGCTGCTCGGGAAGACTGGAGT 1820  
Db 1759 TGCAACATGGGCTGAGGCTGTTGGGGGAGCAACCGCTGCTCGGGAAGACTGGAGT 1818  
Qy 1821 GTACTTTCAAGGACGGTGGGGCAGAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGC 1880  
Db 1819 GTACTTTCAAGGACGGTGGGGCAGAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGC 1878  
Qy 1881 TGTGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGTCATTTGGGCTCGGGAACGC 1940  
Db 1879 TGTGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGTCATTTGGGCTCGGGAACGC 1938  
Qy 1941 TTCTACAGATATGAAAAATTTTGGCTCGATGATGTTTCTCTGTGATGGAAGTACAGTCAGA 2000  
Db 1939 TTCTACAGATATGAAAAATTTTGGCTCGATGATGTTTCTCTGTGATGGAAGTACAGTCAGA 1998  
Qy 2001 TCTCTGGTCATGCAAGAACAGTGGTGGGGAATTAATGACTGCAGTCAAGTGAAGATG 2060  
Db 1999 TCTCTGGTCATGCAAGAACAGTGGTGGGGAATTAATGACTGCAGTCAAGTGAAGATG 2058  
Qy 2061 TGGAGTGTCTGTTCTGATGATCGGATGAGCTGAGGCTTGTGGTGGAGCAGCAG 2120  
Db 2059 TGGAGTGTCTGTTCTGATGATCGGATGAGCTGAGGCTTGTGGTGGAGCAGCAG 2118  
Qy 2121 GTGTCTGGAAAAATTTGAGGTGAATGTCCAGGCTGCCGTGGGAATTTCTGTCTAATGG 2180  
Db 2119 GTGTCTGGAAAAATTTGAGGTGAATGTCCAGGCTGCCGTGGGAATTTCTGTCTAATGG 2178  
Qy 2181 CTGGGAATGAACATTTGCTGAAATTTTGGCAGGCACTTGAATGTGGTGTGCAATCAG 2240  
Db 2179 CTGGGAATGAACATTTGCTGAAATTTTGGCAGGCACTTGAATGTGGTGTGCAATCAG 2238  
Qy 2241 GGTCTCCAGAGAGCCCTCATTTCCAGAAAGAACATTAACATCTTAATCTCGAATCTCG 2300  
Db 2239 GGTCTCCAGAGAGCCCTCATTTCCAGAAAGAACATTAACATCTTAATCTCGAATCTCG 2298  
Qy 2301 CTGCACTGGAGGGAAGCCCTCTCTCTGGGATTTGTATACGATGGGATGGAACACAGCTGC 2360  
Db 2299 CTGCACTGGAGGGAAGCCCTCTCTCTGGGATTTGTATACGATGGGATGGAACACAGCTGC 2358  
Qy 2361 GTGTCTAATTAATATGGAAGCAAGTTTGTATGCTCAGCCCAAGCAGCAGCAGCTGCT 2420  
Db 2359 GTGTCTAATTAATATGGAAGCAAGTTTGTATGCTCAGCCCAAGCAGCAGCAGCTGCT 2418  
Qy 2421 TGGAGCTGATGCCCCCTGCTCTGGAGCAAGTTTGTATGCTCAGCCCAAGCAGCAGCTGCT 2480  
Db 2419 TGGAGCTGATGCCCCCTGCTCTGGAGCAAGTTTGTATGCTCAGCCCAAGCAGCAGCTGCT 2478  
Qy 2481 TGTCTGTGATTTCT 2540  
Db 2479 TGTCTGTGATTTCT 2538  
Qy 2541 TGGAGATGCCATATCTCTTTCTGTGGAGATCATTGGAAGGGAATGCTCTAATCTG 2600  
Db 2539 TGGAGATGCCATATCTCTTTCTGTGGAGATCATTGGAAGGGAATGCTCTAATCTG 2598  
Qy 2601 GSCCGAAAAATTCAGTGTGAAGGAGTGAACACTCACCTTGCATTTATGCCCATTTCTCA 2660  
Db 2599 GSCCGAAAAATTCAGTGTGAAGGAGTGAACACTCACCTTGCATTTATGCCCATTTCTCA 2658  
Qy 2661 ACATCCGGAAGACACTTGTATCCAGCAGAGAAAGTTGGAGTTGTCTGTCTCCGATATAC 2720  
Db 2659 ACATCCGGAAGACACTTGTATCCAGCAGAGAAAGTTGGAGTTGTCTGTCTCCGATATAC 2718  
Qy 2721 AGATGTCGACTTGTGAATGGAATTCACAGTGTGACGGGCAAGTGGAGTCAACGTGCT 2780  
Db 2719 AGATGTCGACTTGTGAATGGAATTCACAGTGTGACGGGCAAGTGGAGTCAACGTGCT 2778  
Qy 2781 TGGACACTGGGCTCAGTGTGTGACACCCACTGGGACCCAGAGATGCCCGTCTCTATG 2840  
Db 2779 TGGACACTGGGCTCAGTGTGTGACACCCACTGGGACCCAGAGATGCCCGTCTCTATG 2838

QY 2841 CAGACAGCTCAGCTGTGGGAGCTCTCTCAACACAGAGGAGAAATATATGGAGAAG 2900  
 DB 2839 CAGACAGCTCAGCTGTGGGAGCTCTCTCAACACAGAGGAGAAATATATGGAGAAG 2898  
 QY 2901 AAGTGTTCGTGTGGGAGCAGAGCTTCTCATTCAGTGTAGGAATGAGTCACTCTGGATAA 2960  
 DB 2899 AAGTGTTCGTGTGGGAGCAGAGCTTCTCATTCAGTGTAGGAATGAGTCACTCTGGATAA 2958  
 QY 2961 CTGTCAAAATGACAGTTCCTTGGAGCACTCCCTGTATCCATGGAATACCTCTGTGTAT 3020  
 DB 2959 CTGTCAAAATGACAGTTCCTTGGAGCACTCCCTGTATCCATGGAATACCTCTGTGTAT 3018  
 QY 3021 CTCACAGGAAGCCTGACCCAGCAGCTGTTCATCCCTCGCAAAATGATCTCAACCCATA 3080  
 DB 3019 CTCACAGGAAGCCTGACCCAGCAGCTGTTCATCCCTCGCAAAATGATCTCAACCCATA 3078  
 QY 3081 TTTGTCTGCAGTTCACAGAGGCGAGTCTTTGTATCTCTTTAGAGGACAAAGCTCGGCT 3140  
 DB 3079 TTTGTCTGCAGTTCACAGAGGCGAGTCTTTGTATCTCTTTAGAGGACAAAGCTCGGCT 3138  
 QY 3141 AGTGGATGGGAGCAGCGCTGTCCGGGAGAGTAGAGATCTATCACGACGGCTTCCTGGGG 3200  
 DB 3139 AGTGGATGGGAGCAGCGCTGTCCGGGAGAGTAGAGATCTATCACGACGGCTTCCTGGGG 3198  
 QY 3201 CACCATCTGTGATCAGCGCTGGAGCCTGTGAGCGATGCCAGTGGTGCTCAAAAGCTGGG 3260  
 DB 3199 CACCATCTGTGATCAGCGCTGGAGCCTGTGAGCGATGCCAGTGGTGCTCAAAAGCTGGG 3258  
 QY 3261 CTGTGGAGTGGCTTCAATGCCAGCGTCTCTGCTCACTTTGGGAGGGGTACAGGCCCAT 3320  
 DB 3259 CTGTGGAGTGGCTTCAATGCCAGCGTCTCTGCTCACTTTGGGAGGGGTACAGGCCCAT 3318  
 QY 3321 CTGCTCGATGACCTGAACCTGCACAGGAACGGAGTCCCACTTTGGCAGTCCCTTCCCG 3380  
 DB 3319 CTGCTCGATGACCTGAACCTGCACAGGAATGGAGTCCCACTTTGGCAGTCCCTTCCCG 3378  
 QY 3381 CGGCTGGGAGCAGCAGCTGCAGGACACAGAGGAGCGAGGGGTATCTCTCAAAAT 3440  
 DB 3379 CGGCTGGGAGCAGCAGCTGCAGGACACAGAGGAGCGAGGGGTATCTCTCAAAAT 3438  
 QY 3441 CACAGCCTTGAGGCTTACAGTGAACCTGAACAGAGAGCTGTCTGGGAGATGGAGT 3500  
 DB 3439 CACAGCCTTGAGGCTTACAGTGAACCTGAACAGAGAGCTGTCTGGGAGATGGAGT 3498  
 QY 3501 CTTCTATAACGGGACCTGGGAGCAGCTCGGAGGAGGAACATCACAGAGCTATAGCAGG 3560  
 DB 3499 CTTCTATAACGGGACCTGGGAGCAGCTCGGAGGAGGAACATCACAGAGCTATAGCAGG 3558  
 QY 3561 CATTGTGTGAGGAGCTGGGAGATGGAGTTGTACAGCTTCGCTTATC 3620  
 DB 3559 CATTGTGTGAGGAGCTGGGAGATGGAGTTGTACAGCTTCGCTTATC 3618  
 QY 3621 TAAGACAGGCTTGTGTATGTGGTGGATGACATTCAGTGTCTTAAACGAGATCTC 3680  
 DB 3619 TAAGACAGGCTTGTGTATGTGGTGGATGACATTCAGTGTCTTAAACGAGATCTC 3678  
 QY 3681 CATATGGCAGTGCCTGTCTGCCCATGGGAGCGAAGATCTCCAGCCGACAGAGAGAC 3740  
 DB 3679 CATATGGCAGTGCCTGTCTGCCCATGGGAGCGAAGATCTCCAGCAGCAGAGAGAC 3738  
 QY 3741 CTGGATACATGTGAAGATGAATAAGAGTCCGTGGAGGAGACACCGAGTCTCTGGAG 3800  
 DB 3739 CTGGATACATGTGAAGATGAATAAGAGTCCGTGGAGGAGACACCGAGTCTCTGGAG 3798  
 QY 3801 AGTGGAGATCTGGCAGCAGGCTCTCTGGGACACAGTGTGATGACCTCTGGGACCTGGC 3860  
 DB 3799 AGTGGAGATCTGGCAGCAGGCTCTCTGGGACACAGTGTGATGACCTCTGGGACCTGGC 3858  
 QY 3861 CGAGCGGAAGTGTGTCTCAGCAGCTGGGCTGTGGCTCTCTCTGGCTGCCCTGAGGGA 3920  
 DB 3859 CGAGCGGAAGTGTGTCTCAGCAGCTGGGCTGTGGCTCTCTCTGGCTGCCCTGAGGGA 3918  
 QY 3921 CGCTTCTGTGGCCAGGGAAGTCTGGTGGATGACATGCGGTCCAAAGGAA 3980

DB 3919 CGCTTCGTTTGGCCAGGGAATCGAACCATCTGGTGTGATGACATGCGGTGCAAGGAA 3978  
 QY 3981 TGAGTCAATTTCTATGGGAGTCTCAGCGCAAAACCCCTGGGAGCAGAGTGAATGTGGACAA 4040  
 DB 3979 TGAGTCAATTTCTATGGGAGTCTCAGCGCAAAACCCCTGGGAGCAGAGTGAATGTGGACAA 4038  
 QY 4041 GGAAGATCTGCGGTGAGGTGCTCTGGAGAGCTGCTGAAATCAGTGAATGCTCTCTCAGG 4100  
 DB 4039 GGAAGATCTGCGGTGAGGTGCTCTGGAGAGCTGCTGAAATCAGTGAATGCTCTCTCAGG 4098  
 QY 4101 TCATTTAGCACTATTTTATCCAGTATCTTTGGGCTCTCTCTCCCTGGTCTCTCTTATCT 4160  
 DB 4099 TCGTTTAGCACTATTTTATCCAGTATCTTTGGGCTCTCTCTCCCTGGTCTCTCTTATCT 4158  
 QY 4161 ATTTCTCAGCTGGTCCGAGTTTCAGAAACAAACATCTGCCCCCTCAGAGTTTCAACAG 4220  
 DB 4159 ATTTCTCAGCTGGTCCGAGTTTCAGAAACAAACATCTGCCCCCTCAGAGTTTCAACAG 4218  
 QY 4221 AAGGAGGGTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCTCAAGAGAG 4280  
 DB 4219 AAGGAGGGTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCTCAAGAGAG 4278  
 QY 4281 GGACCCACATGGCAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAG 4340  
 DB 4279 GGACCCACATGGCAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAG 4338  
 QY 4341 CGACACATCGCTGTGGGAGTTCTTCCCTGCTCTGAAGCCCAAAATGACTTTAGACTTC 4400  
 DB 4339 CGACACATCGCTGTGGGAGTTCTTCCCTGCTCTGAAGCCCAAAATGACTTTAGACTTC 4398  
 QY 4401 CAGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAAACACTTTTAAATGAAT 4460  
 DB 4399 CAGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAGACAAACACTTTTAAATGAAT 4458  
 QY 4461 AAAGAGCAAGTCAAGTTGCCCTATGAAAACCTTTGCCAATTAACATTTCTGAACAAATAG 4520  
 DB 4459 AAAGAGCAAGTCAAGTTGCCCTATGAAAACCTTTGCCAATTAACATTTCTGAACAAATAG 4518  
 QY 4521 GAGAAGCAGCTAAATGATAAAGACTGGTGAATAATAAAATTTGAATTTATATATCAGTGT 4590  
 DB 4519 GAGAAGCAGCTAAATGATAAAGACTGGTGAATAATAAAATTTGAATTTATATATCAGTGT 4578  
 QY 4581 T 4581  
 DB 4579 T 4579

RESULT 2  
 AX133829  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

AX133829 3670 bp DNA linear PAT 15-MAY-2001  
 Sequence 15 from Patent WO0119856.  
 AX133829  
 AX133829.1 GI:14139781

human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 3670)  
 Shimkets, R.A., Fernandes, E., Herrmann, J.L., Liu, X., Yang, M. and  
 Boldog, F.L.  
 Secreted human proteins, polynucleotides encoding them and methods  
 of using the same  
 Patent: WO 0119856-A 15 22-MAR-2001;  
 Curagen Corporation (US)  
 Location/Qualifiers  
 1..3670  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 113..3475  
 /note="unnamed protein product"  
 /codon\_start=1

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 source  
 CDS

/protein\_id="CAC38965.1"  
/db\_xref="GI:14139782"  
/translation="MKNDADVQKIRLENRPNSNENRVPTRHKMDGADLELRADGNSNN  
ISGRVEINSHEDWTTICDONKNEQALVCKQGCPESEVFSRRKPSNEARDLWINS  
CSTGVEINSHEDWTTICDONKNEQALVCKQGCPESEVFSRRKPSNEARDLWINS  
QGEWGTVCHDRWRSTNAAYVCKQGCPESEVFSRRKPSNEARDLWINS  
IWDCEHSGWGHNCVHREDVITVTCSDGATWGLRVGSHNRCGRLEVYFQGRWGTVC  
DWNKAAAVVCSQDLCPSIIIGMLGNASTGYK IWLDDYSCDSDSLWSCNRSW  
GNDCHSESDGVYICSDASOMELRLVGGSSRCAGVEVNOGAVGILCANGWCHLNAE  
VNRQLCEGSAIRVSRPHPTERTLHILMSNSGCTGGEASLWDCIRWEMQDACHLNAE  
EASLISAHROPRLVGLADMPGSRGVKHDATWRSVCDSESLHAANVLCRELNGDA  
ISLVGDFHGFQWAEKFCQEGSETHLALCPVQHPEDTCHSREVGVVCSRYTD  
RLVNGSCQDQVEINVLGHWSLCTDHPEDARVLCRLSCGTALVSTGGYKIGE  
RSVRGHRHFCLENSLNDNCOMTVLGAPPCIHGNTVSVICTSLQPLPCLANVS  
DPYLSAVPEFSALICLEDKRLRLVDGSRACGRVEIYHDFEWGTICDDGMDLSLHVV  
CKLCCGVAFNATVSAHFEGSGPIWLDDNCTGMSHLMWCPKSRGQDCHRIKEDA  
GVICSEETARLYSETETGSCAGLEVYNTGWSVGRNITTAAGVLCVRLCGGEN  
GVSLAPLSKTSQPMVDDIQCPKTHISIMQCLSAPEWRISSPAEETWICEDRIR  
VRGDETCGRVEINSHAGSMGTVCDDSDMLAEVQQLGCGSALALRADSFQGT  
GTIWLDDMRCKGNESFLWDCHEKAPGQSDCHGEDAGVRCQSGLKSLNASGHLALI  
LSSTFGLLLVLFILFTWCRVQKHLPLRVSTRRRGSLLENLFHEMETCLKREDPH  
GTRTSDPTPNHGCEADSDTSLGLVLPASEAPK"  
BASE COUNT 947 a 752 c 1073 g 898 t  
ORIGIN

Query Match 74.8%; Score 3459.8; DB 6; Length 3670;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3464; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 1114 GATGAGCAGATTTGGACTCGGACTAGCAGATGGAAGTAACAAATTTTCAGGGAGAGTA 1173  
DB 200 GATGAGCAGATTTGGAACTGGCACTAGCAGATGGAAGTAACAAATTTTCAGGAAGAGTA 259  
QY 1174 GAGGTGAGAAATTCATGAACAGTGGTGGCAATATCTGACCAAGCTGGAAGAAATGAACAA 1233  
DB 260 GAGGTGAGAAATTCATGAACAGTGGTGGCAATATCTGACCAAGCTGGAAGAAATGAACAA 319  
QY 1234 GCCCTGTGGTGTGTAAGCAGCTAGGATGTCGGTTCAGCGCTTTTGGCAGTCTGCTGCT 1293  
DB 320 GCCCTGTGGTGTGTAAGCAGCTAGGATGTCGGTTCAGCGCTTTTGGCAGTCTGCTGCT 379  
QY 1294 AAACCTAGTAATGAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAG 1353  
DB 380 AAACCTAGTAATGAGCTAGAGACATTTGGATAAACAGCATATCTTGCACTGGGAATGAG 439  
QY 1354 TCAGCTCTCTGGGACTGCACATATCATGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1413  
DB 440 TCAGCTCTCTGGGACTGCACATATCATGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 499  
QY 1414 GATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCTGGGCTCAT 1473  
DB 500 GATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCTGGGCTCAT 559  
QY 1474 AGCCCTGTATGAGGAGATTTGAGGTGAAATACCAAGGAGAGTGGGGAGTGTGTGTCAT 1533  
DB 560 AGCCCTGTATGAGGAGATTTGAGGTGAAATACCAAGGAGAGTGGGGAGTGTGTGTCAT 619  
QY 1534 GACAGATGGACCAAGGAATGCACTGTTGTGTGTAACATTTGGGATGTGGAAGGCT 1593  
DB 620 GACAGATGGACCAAGGAATGCACTGTTGTGTGTAACATTTGGGATGTGGAAGGCT 679  
QY 1594 ATGCAATGTTGTTGATGACCTATTTTAAAGAAAGCATCAGGACCTATTGGCTGGATGAC 1653  
DB 680 ATGCAATGTTGTTGATGACCTATTTTAAAGAAAGCATCAGGACCTATTGGCTGGATGAC 739  
QY 1654 GTTCTTCGATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGAAAG 1713  
DB 740 GTTCTTCGATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGAAAG 799  
QY 1714 CATAATTTGTACACAGAGAGGATGTGATTTAACTGCTCAGGTGATGCCACATGSGGC 1773  
DB 800 CATAATTTGTACACAGAGAGGATGTGATTTAACTGCTCAGGTGATGCCACATGSGGC 859

QY 1774 CTGAGGCTGGTGGGGCGGACAGCAACCGCTGCTCGGGAAGACTGAGAGGTGTACTTTCAAGGA 1833  
DB 860 CTGAGGCTGGTGGGGCGGACAGCAACCGCTGCTCGGGAAGACTGAGAGGTGTACTTTCAAGGA 919  
QY 1834 CGGTGGGCGACAGTGTGTGATGACGGCTGGACACAGTAAGCTGCACCTGTGTGTGTAGC 1893  
DB 920 CGGTGGGCGACAGTGTGTGATGACGGCTGGACACAGTAAGCTGCACCTGTGTGTGTAGC 979  
QY 1894 CAGCTGGACTGCCCATCTTCTATCTATCTGATGAGTGGTCTGGGAAACGCTTCTACAGGATAT 1953  
DB 980 CAGCTGGACTGCCCATCTTCTATCTATCTGATGAGTGGTCTGGGAAACGCTTCTACAGGATAT 1039  
QY 1954 GGAATAATTTGGCTCGATGATGATTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCTATGC 2013  
DB 1040 GGAATAATTTGGCTCGATGATGATTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCTATGC 1099  
QY 2014 AGGAACAGTGGTGGGGAATAATGACTGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2073  
DB 1100 AGGAACAGTGGTGGGGAATAATGACTGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1159  
QY 2074 TCTGATGATCGATCGATATGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2133  
DB 1160 TCTGATGATCGATCGATATGAGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1219  
QY 2134 GTTGAAGTGAATGCTCCAGGCTCCGCTGGGAATCTGTGTGCTAATGGCTGGGGAATGAAC 2193  
DB 1220 GTTGAAGTGAATGCTCCAGGCTCCGCTGGGAATCTGTGTGCTAATGGCTGGGGAATGAAC 1279  
QY 2194 ATTGCTGAAGTGTGTTGACAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGAGAG 2253  
DB 1280 ATTGCTGAAGTGTGTTGACAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGAGAG 1339  
QY 2254 CTTCAATTTTACAGAAAGAACATTTACACATCTTAATGCTCAATTTCTGGCTGCACTGGAGG 2313  
DB 1340 CTTCAATTTTACAGAAAGAACATTTACACATCTTAATGCTCAATTTCTGGCTGCACTGGAGG 1399  
QY 2314 GAAGCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACACTGCTGTCTATTTAAAT 2373  
DB 1400 GAAGCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACACTGCTGTCTATTTAAAT 1459  
QY 2374 ATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 2433  
DB 1460 ATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1519  
QY 2434 CCTGCTCTGAGCTGTTTGAAGTGAACATGACACACATGCGCTCTCTGTGTGTCT 2493  
DB 1520 CCTGCTCTGAGCTGTTTGAAGTGAACATGACACACATGCGCTCTCTGTGTGTCT 1579  
QY 2494 GATTTCTCTCTTCTGCTGCCAATGTGTGTCAGAGAAATTAATTTGTCAGAGATGCCATA 2553  
DB 1580 GATTTCTCTCTTCTGCTGCCAATGTGTGTCAGAGAAATTAATTTGTCAGAGATGCCATA 1639  
QY 2554 TCTCTTTCTGTCGGGAGATCAGCTTTTGGAAAGGAATGGTCTTAACCTTGGCCGCGGAAAGTTC 2613  
DB 1640 TCTCTTTCTGTCGGGAGATCAGCTTTTGGAAAGGAATGGTCTTAACCTTGGCCGCGGAAAGTTC 1699  
QY 2614 CAGCTGGAAGGAGTGAACACTCAGCTTGCATTTATGCCCATTTGTTCAACATCCGGAAGAC 2673  
DB 1700 CAGCTGGAAGGAGTGAACACTCAGCTTGCATTTATGCCCATTTGTTCAACATCCGGAAGAC 1759  
QY 2674 ACTTCTATCCACAGCAGAGAACTTGGAGTGTGTCTGTTCCCGATATACAGATGTCGACAT 2733  
DB 1760 ACTTCTATCCACAGCAGAGAACTTGGAGTGTGTCTGTTCCCGATATACAGATGTCGACAT 1819  
QY 2734 GTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAAGCTTGGACACTTGGGGC 2793  
DB 1820 GTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAAGCTTGGACACTTGGGGC 1879  
QY 2794 TCACCTGTGTACACCACTGGGACCCAGAGATGCCGTGTTCTATGACAGACAGCTCAGC 2853  
DB 1880 TCACCTGTGTACACCACTGGGACCCAGAGATGCCGTGTTCTATGACAGACAGCTCAGC 1919  
QY 2854 TGTGGAGTGTCTCTCAACACAGGAGGAAATATATTGGAGAAAGAGTGTCTGTGTG 2913

3020	Db	CAGGGAACTGGAAACCATCTGGTTTGGATGACATGCGGTGCAAGAGGAATGAGTCATTTCTA	3079
3994	QY	TGGGACTGTCAAGCCCAAAACCCCTGGGGACAGAGTGACTGTGGACACAAGAAAGATGCTGGC	4053
3080	Db	TGGGACTGTCAAGCCCAAAACCCCTGGGGACAGAGTGACTGTGGACACAAGAAAGATGCTGGC	3139
4054	QY	GTGAGGTGCTCTGGACAGTCGCTGAAATCAGTAATGAGTGCCTCTCAGGTCAATTAGCACTT	4113
3140	Db	GTGAGGTGCTCTGGACAGTCGCTGAAATCAGTAATGAGTGCCTCTCAGGTCAATTAGCACTT	3199
4114	QY	ATTTTATCCAGTATCTTTTGGGTCCTCTCCCTCGGTTCTGTTTATTTCTATTTCTCACGTGG	4173
3200	Db	ATTTTATCCAGTATCTTTTGGGTCCTCTCTCCCTGGTTCTGTTTATTTCTATTTCTCACGTGG	3259
4174	QY	TGCCGAGTTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACCAAGAGGAGGGTTCT	4233
3260	Db	TGCCGAGTTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACCAAGAGGAGGGTTCT	3319
4234	QY	CTCGAGGAGAAATTTATTTCCATGAGATGGAGACCTGCCTCAAGAGAGAGACCCACATGGG	4291
3320	Db	CTCGAGGAGAAATTTATTTCCATGAGATGGAGACCTGCCTCAAGAGAGAGACCCACATGGG	3379
4294	QY	ACAAGAACTCTCAGATGACACCCCAACCATGGTTGTGAAGATGCTTAGGCACACATCGCTG	4353
3380	Db	ACAAGAACTCTCAGATGACACCCCAACCATGGTTGTGAAGATGCTTAGGCACACATCGCTG	3439
4354	QY	TTGGGAGTTCTTCTCGCTCTCGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAG	4413
3440	Db	TTGGGAGTTCTTCTCGCTCTCGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAG	3499
4414	QY	ATCAACTCTAAATATCTTTTGAAGGAGACACAACCTTTTAAATGATTAAGAGGAGTCA	4473
3500	Db	CTCAACTCTAAATATCTTTTGAAGGAGACACAACCTTTTAAATGATTAAGAGGAGTCA	3559
4474	QY	AGTTGCCCTATGGAACACTTGTCCAAATTAACATTTCTTGAAACAATAGGAACACGCTAAA	4533
3560	Db	AGTTGCCCTATGGAACACTTGTCCAAATTAACATTTCTTGAAACAATAGGAACACGCTAAA	3619
4534	QY	TTGATAAGACTGGTGATAATAAAAAATTGAATTTATGTATATCAGCTGTATAA	4594
3620	Db	TTGATAAGACTGGTGATAATAAAAAATTGAATTTATGTATATCAGCTGTATAA	3670

precursor.  
AK058038  
ACCESSION  
AK058038.1 GI:16554040  
VERSION  
oligo capping; fis (full insert sequence).  
KEYWORDS  
Homo sapiens synovial membrane (knee) cDNA to mRNA; clone\_lib1 SYN  
SOURCE  
clone:SYN00985.

REFERENCE  
AUTHORS

1. Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S., Ono, Y., Hokuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M., Yamashita, H., Chiba, Y., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T., and Sugano, S.

Sugano, Y.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2249)  
Sugano, Y. and Suzuki, Y.  
Direct Submission  
Submitted (24-OCT-2001) Sumio Sugano, Institute of Medical Science  
University of Tokyo, Laboratory of Genome Structure, Human Genome  
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan

(E-mail: cdnals@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,  
Fax: 81-3-5449-5416)

COMMENT  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan: cDNA full insert sequencing:  
Research Association for Biotechnology (RAB): cDNA library  
construction and 5'-end one pass sequencing: Institute of Medical  
Science, University of Tokyo, Laboratory of Genome Structure, Human  
Genome Center; 3'-end one pass sequencing: RAB; clone selection for  
full insert sequencing: RAB and Helix Research Institute.

FEATURES  
Location/Qualifiers

1..2249

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="SYN00985"

/tissue\_type="synovial membrane (knee)"

/clone\_lib="SYN"

/note="cloning vector: pME18SPL3"

BASE COUNT 559 a 512 c 638 g 540 t

ORIGIN

Query Match 48.5%; Score 2244.2; DB 9; Length 2249;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2333 GTATACGATGGAGTGGAAACAGACTGCGTGTCTATTTAAATATGGAAGCAAGTTGATCT 2392  
Db 1 GTATACGATGGAGTGGAAACAGACTGCGTGTCTATTTAAATATGGAAGCAAGTTGATCT 60  
Qy 2393 GCTCAGCCACAGGAGCCAGGCTGGTGGAGCTGATATGCCCTGCTCTGGAGCTGTTG 2452  
Db 61 GCTCAGCCACAGGAGCCAGGCTGGTGGAGCTGATATGCCCTGCTCTGGAGCTGTTG 120  
Qy 2453 AAGTGAACATCGACACATGCGCTCTGTGTGATTCGATTCCTCTCTTCATGCTG 2512  
Db 121 AAGTGAACATCGACACATGCGCTCTGTGTGATTCGATTCCTCTCTTCATGCTG 180  
Qy 2513 CCAATGTGCTGTCAGAGAAATTAATTTGTGGAGATGCCATATCTCTTTCTGTGGAGATC 2572  
Db 181 CCAATGTGCTGTCAGAGAAATTAATTTGTGGAGATGCCATATCTCTTTCTGTGGAGATC 240  
Qy 2573 ACTTTGGAAGGAATGGTCTACTTGGGCGGAAAGTTCCAGTGTGAAGGGAGTGAA 2632  
Db 241 ACTTTGGAAGGAATGGTCTACTTGGGCGGAAAGTTCCAGTGTGAAGGGAGTGAA 300  
Qy 2633 CTCACCTTGCAATATGCCCCATTTGTTCAACATCCGGAAGACATTTGATCCACAGCAG 2692  
Db 301 CTCACCTTGCAATATGCCCCATTTGTTCAACATCCGGAAGACATTTGATCCACAGCAG 360  
Qy 2693 AAGTTGGAGTTGTCTGTTCCGATATACAGATGTCGACATTTGTAATGGCAAAATCCCACT 2752  
Db 361 AAGTTGGAGTTGTCTGTTCCGATATACAGATGTCGACATTTGTAATGGCAAAATCCCACT 420  
Qy 2753 GTGAGGGCAAGTGGAGATCAACGTGCTTGACACATGGGGCTCACTGTGTACACCCCACT 2812  
Db 421 GTGAGGGCAAGTGGAGATCAACGTGCTTGACACATGGGGCTCACTGTGTGTACACCCCACT 480  
Qy 2813 GGGACCCAGAGATGCCCGTGTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAA 2872  
Db 481 GGGACCCAGAGATGCCCGTGTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAA 540  
Qy 2873 CCACAGGAGAAATATATTTGGAGAAAGTGTGCTGTGGGAGACAGGTTTCATT 2932  
Db 541 CCACAGGAGAAATATATTTGGAGAAAGTGTGCTGTGGGAGACAGGTTTCATT 600  
Qy 2933 GCTTAGGGAATGAGTCACCTTCTGATTAATCTCAATGACAGTCTTGGAGCACTCCCT 2992  
Db 601 GCTTAGGGAATGAGTCACCTTCTGATTAATCTCAATGACAGTCTTGGAGCACTCCCT 660  
Qy 2993 GTATCCATGGAATACTGCTCTGTGTATCTGCACAGGAAGCTGACCCAGCCACTGTTTC 3052  
Db 661 GTATCCATGGAATACTGCTCTGTGTATCTGCACAGGAAGCTGACCCAGCCACTGTTTC 720  
Qy 3053 CATGCCTCGCAATGTATCTGACCCATATTTGCTGCAAGTTCCAGAGGGGAGTGTCTTGA 3112

Db 721 CATGCTCGCAAAATGATATCTGACCCATATTTGCTGCAAGTCTCCAGAGGGCAGTGTCTGA 780  
Qy 3113 TCTCTTTAGAGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTCTCCCGGAGAG 3172  
Db 781 TCTCTTTAGAGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTCTCCCGGAGAG 840  
Qy 3173 TAGAGATCTATCAGACAGCGCTTCTGGGGCAGCATCTGTGATGACGCTGGGACCTGAGCG 3232  
Db 841 TAGAGATCTATCAGACAGCGCTTCTGGGGCAGCATCTGTGATGACGCTGGGACCTGAGCG 900  
Qy 3233 ATGCCACGCTGGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCAGGCTCTCTG 3292  
Db 901 ATGCCACGCTGGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCTTCAATGCCAGGCTCTCTG 960  
Qy 3293 CTCACCTTTGGGAGGGGTGTCAGGCGCCATCTGGCTGGATGACCTGAACCTGCACAGAAATGG 3352  
Db 961 CTCACCTTTGGGAGGGGTGTCAGGCGCCATCTGGCTGGATGACCTGAACCTGCACAGAAATGG 1020  
Qy 3353 AGTCCCACTTGTGCACTGCGCTTCCCGCGCTGGGGCAGCAGCAGCTGCAGGCAACAAG 3412  
Db 1021 AGTCCCACTTGTGCACTGCGCTTCCCGCGCTGGGGCAGCAGCAGCTGCAGGCAACAAG 1080  
Qy 3413 AGGACGCAAGGGTCACTGTCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAAA 3472  
Db 1081 AGGACGCAAGGGTCACTGTCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAAA 1140  
Qy 3473 CAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGCGAGCTGCGCA 3532  
Db 1141 CAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGGACCTGGGCGAGCTGCGCA 1200  
Qy 3533 GGAGGAACATCACCACAGCCATAGCAGCATTTGTGTGAGGAGCTGGGCTGTGGGGAGA 3592  
Db 1201 GGAGGAACATCACCACAGCCATAGCAGCATTTGTGTGAGGAGCTGGGCTGTGGGGAGA 1260  
Qy 3593 ATGAGTTGTGAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTTCATGTGGTGGATG 3652  
Db 1261 ATGAGTTGTGAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTTCATGTGGTGGATG 1320  
Qy 3653 ACATTCAGTGTCTTAAACGCATATCTCCATATGCACTGCTGCTCCCATCGGAGC 3712  
Db 1321 ACATTCAGTGTCTTAAACGCATATCTCCATATGCACTGCTGCTCCCATCGGAGC 1380  
Qy 3713 GAAGAAATCTCCAGCCAGCAGAAAGACCTGGATCACATGTGAAGATAGAATAAGAGTGC 3772  
Db 1381 GAAGAAATCTCCAGCCAGCAGAAAGACCTGGATCACATGTGAAGATAGAATAAGAGTGC 1440  
Qy 3773 GTGGAGGAGACACCGAGTGTCTTGGGAGTGGAGATCTGGCAGCAGGCTCTCTGGGCA 3832  
Db 1441 GTGGAGGAGACACCGAGTGTCTTGGGAGTGGAGATCTGGCAGCAGGCTCTCTGGGCA 1500  
Qy 3833 CAGTGTGTGATGACTCTCTGGGACCTGGCGGAGGAGTGTGTGTGAGAGCTGGGCT 3892  
Db 1501 CAGTGTGTGATGACTCTCTGGGACCTGGCGGAGGAGTGTGTGTGAGAGCTGGGCT 1560  
Qy 3893 GTGCTCTGCTCTGCTGCTGCTGAGGAGCGCTTCTTTGGCCAGGGAACCTGGAACCATCT 3952  
Db 1561 GTGCTCTGCTCTGCTGCTGCTGAGGAGCGCTTCTTTGGCCAGGGAACCTGGAACCATCT 1620  
Qy 3953 GTTGTGATGACATGCGGTGCAAGAAATGAGTCAATTTCTATGGAGCTGTACGCCAAGC 4012  
Db 1621 GTTGTGATGACATGCGGTGCAAGAAATGAGTCAATTTCTATGGAGCTGTACGCCAAGC 1680  
Qy 4013 CCTGGGACAGAGTCACTGTGGACACAAGGAAGTGTGTGGGCTGAGGCTGTGTGAGCAGT 4072  
Db 1681 CCTGGGACAGAGTCACTGTGGACACAAGGAAGTGTGTGGGCTGAGGCTGTGTGAGCAGT 1740  
Qy 4073 CGCTGAATCACTGAATGCCCTCCTCAGTCAATTTAGCACTTATTTATCCAGTATCTTTG 4132  
Db 1741 CGCTGAATCACTGAATGCCCTCCTCAGTCAATTTAGCACTTATTTATCCAGTATCTTTG 1800  
Qy 4133 GGCTCCCTTCTCCTGGTCTGTTTATTTCTACCTGTGGCGAGTGTCAACAACAA 4192

Db 1801 GGCTCTCTCTCGTTCTGTTTATTCTATTCTACGTGGTCCGAGTTCAGAAACAA 1860

QY 4193 AACATCGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGAGGAGAAATTTATTC 4252

Db 1861 AACATCGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGGAGAAATTTATTC 1920

QY 4253 ATGAGATGGAGACCTGCCTCAAGAGAGAGGCCACATCGGACAGAACTCAGATGACA 4312

Db 1921 ATGAGATGGAGACCTGCCTCAAGAGAGAGGCCACATCGGACAGAACTCAGATGACA 1980

QY 4313 CCCCCAACCATGTTGTGAAGATGCTAGGACACATCGTGTGGAGGTTCTCTCTGCT 4372

Db 1981 CCCCCAACCATGTTGTGAAGATGCTAGGACACATCGTGTGGAGGTTCTCTCTGCT 2040

QY 4373 CTGAAGCCCAAAATGACATTTAGCTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 4432

Db 2041 CTGAAGCCCAAAATGACATTTAGCTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 2100

QY 4433 TGAAGGAGACAACTTTTAAATGAATAAGAGAGTCAAGTTCCTCTATGGAAACT 4492

Db 2101 TGAAGGAGACAACTTTTAAATGAATAAGAGAGTCAAGTTCCTCTATGGAAACT 2160

QY 4493 TGTCCTAAATTAACATTTCTGAACAAATAGGAGACACTTAAATGATAAGACTGGTGATA 4552

Db 2161 TGTCCTAAATTAACATTTCTGAACAAATAGGAGACACTTAAATGATAAGACTGGTGATA 2220

QY 4553 ATAAATTAACATTTATGATATCACTGTT 4581

Db 2221 ATAAATTAACATTTATGATATCACTGTT 2249

RESULT 4

HSM130AC1 3786 bp mRNA linear PRI 22-NOV-1996

LOCUS H sapiens mRNA for M130 antigen cytoplasmic variant 1

DEFINITION Z22969

ACCESSION Z22969.1 GI:312143

VERSION antigen; antigen M130.

KEYWORDS Homo sapiens.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3786)

AUTHORS Micklem K.K.

TITLE Direct Submission

JOURNAL Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department of Pathology, University of Oxford, Level 1, Maternity Block, John Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom

REFERENCE 2 (bases 1 to 3786)

AUTHORS Law, S.K., Micklem K.J., Shaw, J.M., Zhang, X.P., Dong, Y., Willis, A.C. and Mason, D.Y.

TITLE A new macrophage differentiation antigen which is a member of the scavenger receptor superfamily

JOURNAL Eur. J. Immunol. 23 (9), 2320-2325 (1993)

MEDLINE 93380506

PUBMED 8370408

FEATURES Location/Qualifiers

1..3786

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_type="Macrophage"

102..3557

/codon\_start=1

/product="M130 antigen cytoplasmic variant 1"

/protein\_id="CAA80542.1"

/db\_xref="GI:312144"

/translation="MVLLDSSGSAFRRHFNVLNPSPTITVLLLSACFTVSSLGGTDK ELRLVDGKNCGRVKEVQEWGTVNCKNGWSMEAVSVICNLQCPALAKPQWANS AGSGRTWMDHVSQGNESALWDCKHDGKNSCTHOOPAGVTCDSGNLEMLTRGG NMGSGRIETKFGRCWCTVCDNENIDHVASVICRQLECGSAVSFSGSSNGEGSPLEW DDLLCNGNESALWNCKHQGWGKNCBDAEDAGVICSKGADLSRLVDGTGSGRLEV RFGEGWGTICDDGWDSDYDAVACKQLGCPATVTAIGRVNASKGFGHILWLDVSCQGH

source

CDS

sig\_peptide 102..221

mat\_peptide 222..3554

/product="M130 antigen cytoplasmic variant 1"

/evidence="experimental"

polyA\_signal 3760..3765

polyA\_signal 3764..3769

polyA\_signal 3768..3773

polyA\_site 3783

BASE COUNT 1006 a 746 c 1074 g 960 t.

ORIGIN

Query Match 31.3%; Score 1450.6; DB 9; Length 3786;

Best Local Similarity 65.4%; Pred. No. 0;

Matches 2182; Conservative 0; Mismatches 1129; Indels 24; Gaps 3;

QY 1002 GCTTGATGTTCTCTCTCGGTAATCAATCTTTCTTTGGGACTGCAGACATTCGG 1051

Db 107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAGACATTTTGTCAACTGAGTCCCTT 1065

QY 1062 AACCTGCAATTTTGAAGTCTTCTCATCAAAACGATGCTGTGATCTGCTCAGATGAGC 1121

Db 167 CACCATTTACTGTGGTCTTACTTCTCAGTCCCTGTTTGTCCACCAGTCTCTTGGAGGAC 2255

QY 1122 AGATTTGGAACCTCGGACTAGCAGATGGAAGTAACTTGTTCAGGAGAGTAGAGGTGAG 1181

Db 227 AGACAAGGAGCTGAGGCTAGTGGTGTGTAACAAACAGTGTAGCGGGAGAGTGAAGTAA 2285

QY 1182 AATTATGACAGCTGTGGACAAATATGTACCAGAACTGGAAGAACTGAACACCCCTTGT 1241

Db 287 AGTCCAGAGGAGTGGGGAACGGTGTGTATAATGCTGGAGCATGGAACGGCTCTCTGT 346

QY 1242 GTTTCTTAAGCAGCTAGGATGCTCCGTTTCAGCGTCTTTGGCAGTCTGCTGCTAAACCTAG 1301

Db 347 GATTGTACCAGCTGGGATGTCCTCACTGCTATCAAGCCCTGGATGGCTAATTCAG 406

QY 1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGACATGGGAATGAGTCACTCT 1361

Db 407 TGCAGGTTCTGGACGATTTGGATGGATCATGTTTCTGTCGTGGAAATGAGTCACTCT 466

QY 1362 CTGGGACTGCACATATGATGGAAACCAAGC---GAACATGCTTCCGAAGATCAGATGC 1418

Db 467 TTGGGATTGCAACATGATGGATGGGGAAGCATAGTAACCTGTAACCAACAGATGC 526

QY 1419 TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGCCTTGTCTGGGCTCATAGCCC 1478

Db 527 TGGAGTCACTGCTCAGATGGATCCAAATTTGGAATGAGGCTCACCGCTGGAGGGAATAT 586

QY 1479 CTGTTATGGAGATTTGGAGTGAATACCAAGAGAGTGGGGAGTGTGTGTGTCATGACAG 1538

Db 587 GTCTTCTGGAAGAAATAGAGATCAAAATTCGAAGCAGCTGGGGAACAGTGTGTGATGATA 646

QY 1539 ATGAGCAGCAAGAAATGACGCTGTTGTGTAAACAAATTTGGGATTTGGGAAGCCTATGCA 1598

Db 647 CTTCAACATAGATCATGCTCTCTGATTTGTAGAACATTTGAATGTGGAAGTCTGTCTAG 706

QY 1599 TGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTTGGTGGATGACCTTTC 1658

Db 707 TTCTCTGGTTCATCTAAATTTTGGAGAGGCTCTGGACCAATCTGTTTGTATGATCTAT 766





Db 647 CTTCAACATAGATCATCTGTCATTTGTGTAGACAACTTGAATGTGGAAGTGTGTCTGAC 706  
Qy 1599 TGTGTTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTGGCTGGAGTACGTTTC 1658  
Db 707 TTTCTCTGGTTCATCTAATTTTGGAGAGGCTCTGGACCAATCTGGTTTGTGATCTTAT 766  
Qy 1659 TTGCATTTGGAAATGAGTCAATATCTGGGACTGTGAACACAGTGGAGTGGGAAAGCATAA 1718  
Db 767 ATGCAACGGAATGAGTCAGCTCTGGAACATCAAGGATGGGGAAGCATAA 826  
Qy 1719 TTGTGTACACAGAGATGTGATTTAACTGCTCAGGTGATGCAACATGGGCGCTGAG 1778  
Db 827 CTGTGATCATGCTGAGGATGCTGAGTGTATTTGCTCAAGGAGCAGACTCTGAGCCTGAG 886  
Qy 1779 GCTGGTGGGCGGAGCAACCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838  
Db 887 ACTGTAGTGSAGTCACTGAATGTTCAAGGAAGATTAGAAGTGAGATTTCAAGGAGATG 946  
Qy 1839 GGGCACAGTGTGTGATGACGCTGGAACAGTAAAGCTGACGCTGCTGTGTGTAGCCAGCT 1898  
Db 947 GGGGACAAATATGTGATGACGCTGGGACAGTTACGATGCTGCTGGCATCAAGCAACT 1006  
Qy 1899 GGACTGGCCATCTCTATCATTTGCGATGGGTCTGGGAACCTTCTACAGGATATGGAA 1958  
Db 1007 GGGATGTCCAACTGCCGTACAGCCATTTGGTGGAGTTAAGCCAGTAAAGGATTTGGACA 1066  
Qy 1959 AATTTGGCTCGATGATTTCTGTGTGAGATGAGTGCAGTCTCTGCTCATGCAAGAA 2018  
Db 1067 CATCTGGCTTGACAGCTTTCTTCCAGGACATGAACCTGCTGCTGGCAATGTAACA 1126  
Qy 2019 CAGTGGGTGGGAAATTAATGACTGCAGTACAGTGAAGATTTGGAGTGTCTGTTCGA 2078  
Db 1127 CCATGAATGGGAAAGCAATTTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 1186  
Qy 2079 TGCATCGGATATGAGCTGAGCTGTGGTGGAGCAGCAGTGTGCTGGAAAGTTGA 2138  
Db 1187 TGGATCAGATCTGAGCTTAAGACTTAGAGTGGAGGCGCCGCTGTGCTGGGACAGTTGA 1246  
Qy 2139 GGTGAATCTCAGGCTGCGTGGGAATTTCTGTGTCTAATGGCTGGGAATGAACATTCG 2198  
Db 1247 GGTGGAGATTCAGAGACTGTGTAGGAAGTGTGTGACAGAGCTGGGACTGGAAGAAGC 1306  
Qy 2199 TGAAGTTGTTGAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGAGACCTCA 2258  
Db 1307 TGATGTGTTTGCAGGCACTGGGATGTGGATGCTGCATCAAAACATCTTATCAAGTGA 1366  
Qy 2259 TTTTCACAGAAAGACATTAACATCTTAATCTCGAATTTCTGCTGCTGCTGCTGAGGGAAGC 2318  
Db 1367 CTCGCAATCCAGGCAACAAACACATGGCTGTTCTAAGTGTGTAACGGAAATGAAC 1426  
Qy 2319 CTCTCTCTGGATTTGATACGATGGGAGTGGAAACAGAGTCTGCTGCTCATTTAAATATGA 2378  
Db 1427 TTTCTTTGGGACTGCAAGAACTGCAATGGGTGGGACTTACCCTGTGATCATATGAAGA 1486  
Qy 2379 AGCAAGTTTGTCTGCTCAGCCACAGGACCCAGGCTGGTTGGAGCTGATATGCCCTG 2438  
Db 1487 AGCCAAAATTAACCTGCTCAGCCACAGGGAACCCAGACTGGTTGGAGGGGACATTTCCCTG 1546  
Qy 2439 CTCTGGAGCTGTGAAGTGAACATGACACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2498  
Db 1547 TTTCTGGAGCTGTGAAGTGAAGCAGTGGTGAACGCTGGGGCTCCATCTGTGATTCGGACTT 1606  
Qy 2499 CTCTCTTCATGCTGCCAATGTGCTGTCAGAGAAATTAATTTGTGGAGATGCCATATCTCT 2558  
Db 1607 CTCTCTGAAGCTGCCAGCTTCTATGACGGGAATTAACGTTGACAGTGGCAGAGTTGCTCTAT 1666  
Qy 2559 TTCTGTGGGAGATCACTTTGGAAAAGGAATGGTCTAACTTTGGGCCGAAAAGTTCCAGTG 2618  
Db 1667 CTTGGGGGAGCTCACTTTGGAGAGGGAATGGACAGATCTGGGCTGAAGATTTCCAGTG 1726  
Qy 2619 TGAAGGAGTGAACACTCACTTGCATTTATGCCCCATTTGTCACATTCGGGAGACACTTG 2678  
Db 1727 TGAGGAGCATGATGCCATCTTCTACTCTGCCAGTAGCAGCCGCCGAGGAAGCACTTG 1786

Qy 2679 TATCCACAGCAGAGAAAGTTGGAGTTGTCTGTCTCCGATATACAGATGTCCGACTTGTGAA 2738  
Db 1787 TAGCCACAGCAGGAGTGTGGAGTAGTCTGCTCAGATATACAGAAATTCGCTTGGTGA 1846  
Qy 2739 TGGCAAAATCCAGTGTGACGGGCAAGTGAGATCAACGTGTGGACACTGGGGCTCACT 2798  
Db 1847 TGGCAAGACCCGCTGTGAGGCGAGAGTGGAGCTCAAAACGCTTGGTGTGGGATCCCT 1906  
Qy 2799 GTGTGACACCCACTGGGACCCAGAGATGCCGCTGTCTATTCACACAGCTCAGCTGCG 2858  
Db 1907 CTGTAACTCTCTACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966  
Qy 2859 GACTGCTCTCTCAACACAGGAGGAAAAATATATTGGAGAAAGAACTGTTCGTGTGGGG 2918  
Db 1967 AGTTGCCCTTTCTACCCCGAGGAGGACGTTTGGAAAAAGGAATGTGCAGATCTGGAG 2026  
Qy 2919 ACAGAGTTTTCATTTAGGAAATGAGTCACTTCTGTAATCTGTAATCTCAATGACAGTCT 2978  
Db 2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCAGATTTGCTGTAACTGCTCT 2086  
Qy 2979 TGGAGCCTCTCCCTGTATCCATGGAATACTGTCTCTGTGATCTGCACAGGAAGCTGAC 3038  
Db 2087 AGTGTCTTCAATGCTCTTACAGCAAGTGGCCTCTGTATCTGCTCAGGAACAGCTC 2146  
Qy 3039 CCAGCCTGTTTCCATGCCCTCGCAAAATGTATCTGACCCATATTTGCTGCAAGTTCACGA 3098  
Db 2147 CCAACACTGCTCTGTCGAATTCATCTGTTTGGGCCCAACAAAGGCTTACCATCCAGA 2206  
Qy 3099 GGGCAGCTCTTGTGATCTCTTAGAGGACAAAGGCTCGGCTAGTGGATGGGAGCAGCG 3158  
Db 2207 AGAAGTCTCTGGGCTGCATAGAGAGTGGTCAACTGCGCTTGGTAAATGGAGAGGTCG 2266  
Qy 3159 CTGTGCGGAGAGTAGAGATCTATCAGCAGGCTTCTGGGGCACATCTGTGATGACGG 3218  
Db 2267 CTGTGCTGGGAGTAGAGATCTATCATGAGGCTCTTGGGGACCTCTGTGATGACAG 2326  
Qy 3219 CTGGACCTGAGCGATGCCAGCTGTGTCTAAAAGCTGGGCTGTGGAGTGGCTTCAA 3278  
Db 2327 CTGGACCTGAGTGTGATGCCAGCTGTTTGCAGACAGCTGGGCTGTGGAGAGGCTTAA 2386  
Qy 3279 TGCACAGTCTCTGCTCACTTTGGGAGGGTGCAGGGCCCATCTGGCTGGATGACCTGA 3338  
Db 2387 TGCCACTGTTCTGCTCATTTTGGGGAAGGAACAGGGGCCCATCTGGCTGGATGAGATGA 2446  
Qy 3339 CTGCACAGAAAGGAGTCCCACTTGTGCGAGTGGCTTCCCGCGCTGGGGGACAGCA 3398  
Db 2447 ATGCAATGGAAGAAGTCCCGCATTTGGCAGTGGCCATTCACAGGCTGGGGGACGAAA 2506  
Qy 3399 CTGCAGGCAACAGGAGGACGAGGGTCACTGTCTCAGAAATTCACAGCTTGGAGCTCTA 3458  
Db 2507 TTGCAGGCAACAGGAGGATCGGGAGTTATCTGCTCAGAAATTCATGCTCTCAGACTGAC 2566  
Qy 3459 CAGTCAAACTGAAACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTATAACGGGACCTG 3518  
Db 2567 CAGTGAAGCCAGCAGAGAGGCTGTGACGGGCTCTGGAAGTTTCTTAAATGGAGCTTG 2626  
Qy 3519 GGGCAGCTGCGGAGGAGGAACATCACACAGCCATAGCAGGATGTGTGCGAGCAGCT 3578  
Db 2627 GGGCACTGTTGGCAAGATGAGCATGCTGCAACCACTGTGGGTGTGGTGTGCGAGCAGCT 2686  
Qy 3579 GGGCTGTGGGAGAAATGGAGTTGTGAGCTCGCCCTTTTATCTAAAGACAGGCTCTGGTTT 3638  
Db 2687 GGGCTGTGACAAAGGAAAAATCAACCTGCTCTTTAGACAAAGGCCATGTCCATTC 2746  
Qy 3639 CATGTGGGTGAGTACATTCAGTGTCTTAAACGAGATATCTCCATATGCACTGTGCTGTC 3698  
Db 2747 CATGTGGGTGAGCAATTTAGTGTCTCAAAAGGAGCTGACACGCTGTGGCAGTGCCCATC 2806  
Qy 3699 TGCCCCATGGGAGCAAGAAATCTCCAGCCACAGCAAGAGACCTGATACATGTGAAGA 3758  
Db 2807 ATCTCCATGGGAGAGAGACTGGCCAGGAGGAGCTGGATACATGTGACAA 2866





Query Match	30.5%	Score 1409.4;	DB 6;	Length 3707;
Best Local Similarity	65.3%	Pred. No. 0;		
Matches 2176; Conservative	0;	Mismatches 1131;	Indels 28;	Gaps 6;

[illegible]

QY 2078 ATGCATCGGATATGGAGCTTGGGTGGGAGAGCAGAGGCTGTGCTGAAAGATTG 2137  
Db 1187 ATGGATCAGATCTGGAGCTTAAGACTTAGAGGTGGAGCGCCCTGTGCTGGACAGTTG 1246  
QY 2138 AGGTGAATCTCAGAGTGCCTGGGAATCTCTGTCTTAATGGCTGGGAAATGAACATTG 2197  
Db 1247 AGGTGGAGATTCAGAGACTGTAGGAAGAGTGTGTACAGAGGCTGGGACTGAAAGAG 1306  
QY 2198 CTGAAGTGTCTTCAGGCAACTTGAATGTGGGTCTCAATCAGAGTCTCCAGAGAGCCCTC 2257  
Db 1307 CTGATGTGTTCAGGACAGCTGGGATGTGGATCTGCATCAAAACATCTTATCAAGTGT 1366  
QY 2258 ATTTCAGAAAGAACATTTACACATCTTAATGTGCAATTTCTGGCTGCATCGAGGGGAG 2317  
Db 1367 ACTCCAAAATCCAGCAACAAACACATGCTGTTCTAAGTAGCTGAACGGAATAA 1426  
QY 2318 CTTCTCTCTGGGATCTATACGATGGAGTGGAAACAGAGCTGGCTGCATTTAAATATGG 2377  
Db 1427 CTTCTCTTTGGGACTGCAAGAACTGGCAATGGGTGGGACTTACCTGTGATCACTATGAAG 1486  
QY 2378 AAGCAAGTTTGATCTGCTCAG--CCACAGGAGCCAGGCTGGTGGAGCTGATATGCC 2435  
Db 1487 AAGCCAAAATTAACCTGCTCAGGCCACAGGGNAACCCACACTGGTGGAGGGGACATTC 1546  
QY 2436 CTGCTCTGACGCTGTGAAGTGAACATCGACACATCGCCGCTCTGTCTGATTCGA 2495  
Db 1547 CTGTTCTGACGCTGTGAAGTGAAGCATGGTGACACGCTGGGCTCCATCTGTGATTCA 1606  
QY 2496 TTTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAAATTAATTTGGAGATGCCATATC 2555  
Db 1607 CTTCTCTCTGGAAGCTGCCAGGCTTCTATGACAGGAATTAACAGTGTGGCACAGTTGCTC 1666  
QY 2556 TCTTTCTGTGGAGATCACTTTGGAAAAGGGAATGCTTAACCTTGGCCGCGAAAAGTTCCA 2615  
Db 1667 TATCTGGGGAGCTCACTTTGGAGAGGGAATGGACAGATCTGGGCTGMAAGAAATCCA 1726  
QY 2616 GTGTGAAGGAGTGAACACTCACTTTGCAATATGCCCCATTTGTTCAACATCCGGAAGAC 2675  
Db 1727 GTGTGAGGACATGAGTCCCATCTTCACTCTGCCAGTAGCACCCGCCAGAGGAAC 1786  
QY 2676 TTGTATCCACAGCAGAGAGTTGGAGTTGCTGTGTTCCCATATACAGATGTCGACTTGT 2735  
Db 1787 TTGTATCCACAGCAGAGGATGTTGGAGTGTCTGCTCAAGATACACAGAAATTCGCTTGT 1846  
QY 2736 GAATGCAAAATCCAGTGTGAGGCAAGTGGAGATCAACGCTCTTGGACACTGGGGCTC 2795  
Db 1847 GAATGCAAGACCCGCTGTGAGGCAAGTGGAGCTCAAAACGCTTGGTGGCTGGGATC 1906  
QY 2796 ACTGTGTACACCCACTGGGACCCAGAGATGCCGCTGTCTATGACAGACGCTCAGCTG 2855  
Db 1907 CTTCTGTAACCTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATG 1966  
QY 2856 TGGGACTGCTCTCAACACAGGAGAAATATATTGGAGAAGAGTGTCTGCTGTG 2915  
Db 1967 TGGAGTGGCCCTTTTACCCAGGAGGAGCAGCTTTTGGAAAAGGAAATGGTCAGATCTG 2026  
QY 2916 GGGACACAGCTTTCATTTAGGGAATGAGTCACTTCTGGATAACTGTCAAAATGACAGT 2975  
Db 2027 GAGGCATATGTTCACTGCACTGGGACTGAGCAGCAGATGGGAGATGTCTCTTAAGTGC 2086  
QY 2976 TCTTGGACACCTTCCCTGTATCCATGGAATATCTGTCTCTGTGATCTGCACAGGAAGCCT 3035  
Db 2087 TCTAGTGTCTATATATGCTCTGAGCAAGTGGGCTCTGTAACTCTGCTCAGGAAACCA 2146  
QY 3036 GACCCAGCCACTTTCATGCTCGCAATGTATCTGACCCNATATTGTCTGACGTTCC 3095  
Db 2147 GTCCCAAAACACTGCTCTGCAATTCATCTGCTTTTGGGCCCAACAGGCCCTACCATTCC 2206  
QY 3096 AGAGGCACTGCTTTCATCTGCTTAGAGCAAAACGGCTCGGCTAGTGGATGGGAGAG 3155  
Db 2207 AGAGAAAGTGTGTGGCTGCTATAGAGAGTGTCAACTTCCCTGGTAAATGGAGAGG 2266  
QY 3156 CCGCTGTCCGGAGAGTAGATCTATCACGACGGCTTCTGGGGCACCATCTGTGATGA 3215

Db 2267 TCCCTGTCTGGAGAGTAGATCTATCATGAGGCTCTCTGGGACCATCTGTGATGA 2326  
QY 3216 CCGCTGGGACCTGAGGATGCCACGCTGTGTCTCAAAAGCTGGCTGTGGAGTGGCCTT 3375  
Db 2327 CAGCTGGGACCTGAGTGTGATGCCACGCTGTGTGTTCAGACAGCTGGGCTGTGGAGAGCCAT 2386  
QY 3276 CAATGCCACGCTCTCTGCTCACTTTTGGGAGGGGTTCAGGGCCCATCTGGCTGTGATGACCT 3335  
Db 2387 TAATGCCACTGTCTCTGCTCATTTTGGGAAAGAACAGGCCCCATCTGGCTGTGATGAGAT 2446  
QY 3336 GAACTGCACAGAACGAGAGTCCACCTGTGTGGCAGTCCCTTCCCGGGCTGTGGGGGAGCA 3395  
Db 2447 GAAATCAATGGAAGAAATCCCGCATTTGGCAGTCCCATTCACAGGCTGGGGGAGCA 2506  
QY 3396 CGACTTCAGGCAACAGGAGGAGCGGCTGCTCTCAGAAATTCACAGCCTTTGAGGCT 3455  
Db 2507 AAATTCAGGCAACAGGAGGAGTGTGCGGAGTATCTCTCAGAAATTCATGCTCTGAGACT 2566  
QY 3456 CTACAGTGAACCTGMAACAGAGAGCTGTCTGCGAGATTGGAAGTCTTCTATAACGGAC 3515  
Db 2567 GACCACTGAAGCCAGCAGAGAGGCTGTGCAGGGCTCTGGAAGTTTTCACATGAGC 2626  
QY 3516 CTGGGCAAGCTGCGGAGGAGGAACATCACACAGCCATAGCAGGATTTGTGTGACAGCA 3575  
Db 2627 TTGGGCACTGTGGCAAGAGTAGCATGTCTGAAACCACTGTGGGTGTGGTGTGACAGCA 2686  
QY 3576 SCTGGGCTGTGGGAGAAATGGAGTTGTGAGCCTCGCCCTTTATCTAAGACAGGCTCTGG 3635  
Db 2687 GTGGGCTGTGCAGACAGAGGAAATCAACCTGCTATCTTAGAAGGCCCATGTCAT 2746  
QY 3636 TTTTATGTGGGTGGATGACATTCAGTGTCTTAAAGCCATATCTCCATAT--GCCAGTGCC 3694  
Db 2747 TCCCATGTGGGTGGACAAATGTTTCAAGTGTCCAAAAGGACCTGCACACCTGTGGCAGTGCC 2806  
QY 3695 TGTCTCCCATGAGGAGGAGAAATCTCAGCCAGCAGAGAGAGCCTGGATCAGATGTG 3754  
Db 2807 CATCATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGAGCCTGGATCAGATGTG 2866  
QY 3755 AGATAGAAATAGAGTGGTGGAGGAGACACCGAGTGTCTGGAGAGTGGAGATCTGGC 3814  
Db 2867 AACAAAGATAGAGCTTCAGAAAGGACCCATCTCTGTTCTGGACGTGTGGAGATCTGGC 2926  
QY 3815 AGCAGGCTCTCTGGGACACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGCGGAGATGG 3874  
Db 2927 ATGGAGTTCCTGGGGACAGTGTGTGATGACTCTTGGGACTTTGGACGATGCTCAGGTG 2986  
QY 3875 TGTGTGACAGCTGGGCTGTGGCTGTCTGCTGCTGCTGCGCTGAGGAGAGCTTCTGTTGGCC 3934  
Db 2987 TGTGTCAACAACTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3046  
QY 3935 AGGAACTGGAACTGCTGTTGGATGACATGGGTGCAAGGAAATGAGTCAATTTCTAT 3994  
Db 3047 AGGGACTGAGCCGATATGGCTCAATGAAGTGAAGTGAAGGAAATGAGTCTTCCTTGT 3106  
QY 3995 GGGACTGTCCAGCAAAACCTCTGGGACAGAGTGTGTGACACAAAGGAAATGCTGGCG 4054  
Db 3107 GGGATTTCTCTCCAGACGCTGGGCGCATAGTGTGAGTGTGGGCAAGGAAAGAGCTGAG 3166  
QY 4055 TGAGTGTCTGTGACAGTCTGCTGAA-----ATCACTGAATGCTCTCT 4096  
Db 3167 TGAATTTGCACAGATATTTTCAAGTGCAGAAACCCACAAAAGGCCAACAGGCTGCTCAT 3226  
QY 4097 CAGGCTATTAGCAGTATTTATCCAGTATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCT 4156  
Db 3227 CCGCTCAGTCATCTCTTTATTTGAGTCGGGATCTCTGGGTTGTTCTGTTGGCCATTTTCG 3286  
QY 4157 TTTCTATTTCTCAGCTGGTCCGAGTTTCAGAAACAAAACATCTGCCCTCTCAGAGTTTCAA 4216  
Db 3287 TCGCATATTTCTTCTGACTTAAACGAGACAGAGACA---GCGGCTTGCAGTTTCTCT 3343  
QY 4217 CCAGAGGAGGGTCTCTCTCAGGAGAAATTTATTCATGAGATGGAGAGCTGCTCTCAAGA 4276

Db 3344 CAAGAGAGAGAACTTAGTCCACCAAAATCAATACCGGGAGATCAATTTCTGCTGAATG 3403

QY 4277 GAGAGGACCCACATGGGACAAAGACCTCAGATGAC 4311

Db 3404 CAGATGATCTGGACCTAATGAATTCCTCAGGAGGC 3438

RESULT 8

HSM130AE 3802 bp mRNA linear PRI 22-NOV-1996

LOCUS H.sapiens mRNA for M130 antigen extracellular variant.

DEFINITION 222971

ACCESSION 222971.1 GI:312147

VERSION antigen; antigen M130.

KEYWORDS

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3802)

Micklem.K.K.

Direct Submission

Submitted (14-JUN-1993) Kingsley K.J. Micklem, Nuffield Department of Pathology, University of Oxford, Level 1, Maturity Block, John Radcliffe, Hospital, Headington, Oxford, OX3 9DU, United Kingdom

2 (bases 1 to 3802)

Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C. and Mason, D.Y.

A new macrophage differentiation antigen which is a member of the scavenger receptor superfamily

JOURNAL Eur. J. Immunol. 23 (9), 2320-2325 (1993)

MEDLINE 93380506

PUBMED 8370408

FEATURES

Location/Qualifiers

1..3802

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_type="Macrophage"

102..3551

/codon\_start=1

/product="M130 antigen extracellular variant"

/protein\_id="CAA80544.1"

/db\_xref="GI:312148"

/db\_xref="SPTREMBL:Q07901"

/translation="MVLLDSGSADPRHFVNLSPEPTIVVLLLSACFCVTSLSIGTDK  
EALVDGKNCGRSEVKEVQVEWGTVCNNGNSAEVVCNQLGCPALKAPEWANS  
AGSGRIEMHDSVCRGNSALWDKHDGKNSCHTQDDAGVCSGDSNLEMLRTGG  
NMCSGRIEIKFQGRWGTVDNFDNIDHSAVICRQLECGSAVFSGSSNFGESGPTWF  
DLICNGNESALWCKHOGKHNCDHAEDAGVICSGADLSRLVLDVGTSCSGLV  
RFQEWGTTICDDGWSYDAAVACKOLGCPYATAGRVNASKGFHILWDSVSCOGHE  
PAWOCKHHEMKHVCNHNEDAGVCSGDSLELRGGRSGCACTVEIQRLLGKV  
CDRWGLKEDVVCRLGCGSALKTSIQYIKQATNWLFLSSNCNETSLWDCKNW  
QWGLTCDHYEAKITCSAHPRLVGGDIPCSGRVEYKHGDTWGSICDSDFSLAAS  
VLCRELQCVTVYILGAFHGGNGQIWAEEFQCEGHSHLSLCPVAPRPGTCSHSR  
DVGVCSSTQKTSITGLSVTVKGTGLGSHSLFLPKCLLPVGTFLVNGKTPCEGRV  
ELKTLGAGSLCNSHWDIEDAHLVQOLKCGVALSTPGGAREFGKNGOLWRFHCTG  
TEQHMDCPVTALGSLCSEQVAVSVCNQSQTLSSNCSSSLGPTRTIPEEAVA  
CIESGLRVNGWCGKAGREYIHGEGWGTICDDSDWSDAHVCRVQLGCGEALNATG  
SAHFGGTPIWLDENKGRSRIWQSHGQCNKHKEDAGVICSEFMSRLTS  
EASREACARLEFVFNAGVTGKSMSETTVGVVCRQLGCKADKINPASPDLKMSI  
PMWNVQCPKGFDTLWOCPSPPWEKRLASPEETWITCDNKLRLQEGPTSCSGRVEI  
WHGGSGTVCDSWDLDDAQQVQOLGCGPALKAPKEAFEGOGTGPILNEVYKQNE  
SLWDCPAREWHSCEGHKEDAAVNCTDLSVQKTPQKATIGRSSRQSFSFIAGVILGV  
LIAIFVAFLEFLTKRQRORLAVSSRGENLHQIYQIREMNSCLNADDLMLNMSGHS  
EPH"

sig\_peptide 102..221

mat\_peptide 222..3548

/product="M130 antigen extracellular variant"

/evidence=experimental

polyA\_signal 3776..3781

polyA\_signal 3780..3785

polyA\_signal 3784..3789

polyA\_site 3799

BASE COUNT 1009 a 751 c 1081 g 961 t

## ORIGIN

Query Match 28.9%; Score 1339.2; DB 9; Length 3802;  
Best Local Similarity 63.5%; Pred. No 0;  
Matches 2179; Conservative 0; Mismatches 1128; Indels 123; Gaps 4;

QY 1002 GCTTGATGGTCTCCTGCTCCGGTAATGAATCTTTTCTTTGGGACTGCGAGATTCCTCGG 1061

Db 107 GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166

QY 1062 AACCGTCAATTTTGCATGCTTTCATCAAAACGATGCTGCTGATCTGCTCAGATCGAGC 1121

Db 167 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTGTCCACCACTTCTCTTGGAGAAC 226

QY 1122 AGATTTGGAACTCGCACTAGCAGATGGAAGTAACAATTTTCAGGGAGAGTAGAGGTGAG 1181

Db 227 AGACAAGGAGCTCAGGCTAGTGGTGTGAAACAGCTGTAGCGGGAGAGTGGAACTGAA 286

QY 1182 AATTCTATGAACAGCTGGTGGACAATATGTACCAGAACTTGGAGAATGAACAAGCCCTTGT 1241

Db 287 AGTCCAGGAGGAGTGGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT 345

QY 1242 GGTTCGTAAGCAGCTAGGATGCTCGGTTTCACGCTGTTTGGCAGTCTGCTCTAAACCTAG 1301

Db 347 GATTGTAAACCACTGGGATGCTCCAACTGCTATCAAGCCCTGGATGGGCTAATTCACG 406

QY 1302 TAATGAAGCTAGAGACATTTGGATAAACAACAGCATATCTTGCACCTGGGAATCAGTCACTCT 1361

Db 407 TGCAGGTTCTGGACGCACTTTGGATGGATCATGTTTCTGTGCTGGGAATCAGTCACTCT 466

QY 1362 CTGGGACTGCACATATGATGGAAGAAAGCAAGC---GAACATGCTTCCGAAGATCAGATGC 1418

Db 467 TTGGGATGTCAAACATGATGGATGGGAAAGCATAGTAACCTGTACTCACCACAAGATGC 1526

QY 1419 TGCAGTAATTTGTTCTGATAAGGCAGATCTGGACCTTAAGGCTTGTGCGGCTCATAGCCC 1478

Db 527 TGGAGTGACCTGCTCAGATGGATCCAAATTTGGAATGAGGCTCAGCGTGGAGGGAATAT 1586

QY 1479 CTCTTATGGAGATTTGGAGTGAATACCAAGAGAGTGGGGAGCTGTGTGTCATGACAG 1538

Db 587 GTCTTCTGGAAGAATAGAGATCAAAATCCAAAGACGCTGGGAAACAGTGTGTGATGATAA 1645

QY 1539 ATGGACACAAAGAAATGCAGCTGTTGTGTAACAAATTTGGGATGTGGAAGCCTATGCA 1598

Db 647 CTTCAACATAGATCATGTCATCTGTCATTTGTAGAACACTTGAATGTGGAAGTCTGTGTCAG 1706

QY 1599 TGTGTTTGGTATCACCTATTTTAAAGAAGCATCAGGACCTATTTTGGTGTGATGACGTTTC 1658

Db 707 TTTCTCTGGTTTCATCTAAATTTTGGAGAAGCTCTGGACCAATCTGGTTGATGATCTTAT 1766

QY 1659 TTGCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAGCATAA 1718

Db 767 ATGCAACGGAATGAGTCACTGCTCTCTGGAACATCAAGACATCAAGGATGGGGAAGCATAA 1826

QY 1719 TTGTGTACACAGAGAGATGATGTTGTAACCTGCTCAGGTGATGCAACATGGGGCCCTGAG 1778

Db 827 CTGTGATCATGCTGAGGATGCTGCGACTGATTTCTCTCAAGGGAGCAGATCTGAGCCCTGAG 1886

QY 1779 GCTGTTGGGCGGAGCAACCCGCTGCTCGGAAGCATGGAAGTCTACTTTTCAAGGACGGTG 1838

Db 887 ACTGGTAGATGGAGTCACTGAATGTTTTCAGGAAGATTAGAAGTGAAGTCCCAAGGAATG 1946

QY 1839 GGACACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTGACGCTGTGTGTGTAGCCAGCT 1898

Db 947 GGGGACAATATGTGATGACGGCTGGACAGTACGATGCTGTGTGTCATGCAAGCAACT 1006

QY 1899 GGACTCCCATCTTCTATCATTTGGCATGGTCTGGGAACGCTTCTACAGGATATGGAA 1958

Db 1007 GGGATGTCCAACCTGCGCTCACAGCCATTTGGTTCAGTAAAGCCAGTAAGGGATTTGGACA 1066

QY 1959 AATTTCGCTCGATGATGTTTTCCTGTGATGGAGATGAGTCAAGATCTCTGCTGTCATGAGGAA 2018

Db 1067 CATCTGCTTGACAGGCTTCTTGGCAGGACATGAACCTGCTGCTGCAATGTAACA 1126

QY 2019 CAGTGGTGGGAAATTAATGACTGCGAGTGCAGTGAAGATGTTGGAGTGCATCTGTTCTGA 2078  
DB 1127 CCATGAATGGGGAAGCAATATTGCAATCAACATGAGATGCTGGCGTGACATGTTCTGA 1186  
QY 2079 TGCATCGGATATGGAGCTGAGCGTTTGTGGTGGAGCAGCAGGTTGTGCTGGGAAAGTTGA 2138  
DB 1187 TGGATCAGATCGGAGCTAAGACTTAGAGTTGGAGCGCCCTGCTGCTGGGACAGTTGA 1246  
QY 2139 GGTGAATGCCAGGTCGCGTGGGAATTTCTGTGCTTAATGGTGGGGAATCAACATGTC 2198  
DB 1247 GGTGAGATTCAGAGACTGTTAGGGAAGGTGTGACAGAGGCTGGGACTGAAAGAAGC 1306  
QY 2199 TGAAGTGTGTCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCCCTCA 2258  
DB 1307 TGATGTGTTGCGAGCAGCTGGGATGGATCTGCATCAAAACATCTTATCAAGTGA 1366  
QY 2259 TTTTCAGAAAGAACATTACACATCTTAATGTCGAATTTGCTGCATCGAGGGGAAGC 2318  
DB 1367 CTCCAAAATCCAGGCAACAAACACATGGCTGTTCTAAGTAGCTTAACGGGAATGAAC 1426  
QY 2319 CTCCTCTGGGATTTATACGATGGAGTGGGAACAGACTGCTGCTCATTTTAAATATGA 2378  
DB 1427 TTCTCTTTGGGACTGCAGAACTGGCAATGGGGTGGACTTACTGTGATCATATGAAGA 1486  
QY 2379 AGCAAGTTTGATCTGCTCAGCCCAAGCAGGCGCCAGGCTGTTGGAGTGTATATGCCCTG 2438  
DB 1487 AGCCAAAATACCTGCTCAGCCCAAGCAGGCAACCCAGACTGGTTGGAGGCAATTCCTG 1546  
QY 2439 CTCCTGAGCTGTTGAAGTGAACATGCAGACACATGCGGCTGCTGCTGTGATCTGATTT 2498  
DB 1547 TTTCTGAGCTGTTGAAGTGAACATGCTGACAGTGGTGCACAGTGGGCTCCATCTGTGATTCGGACTT 1606  
QY 2499 CTCCTCTGATGCTGCAATGCTGTCGAGAGAAATTAATTTGAGAGTGCATATCTCT 2558  
DB 1607 CTCCTGGAAGCTGCCAGCTTCTATGAGGGAATTAACAGTGTGCACAGTTGCTCTCTAT 1666  
QY 2559 TTCTGTGGGAGATCACTTTGAAAAGGAATGGTCTAATCTGGGCGAAAGTTCCAGTG 2618  
DB 1667 CCTGGGGGAGCTCACTTTGGAGAGGAAATGGACAGATCTGGGCTGAAGATTTCCAGTG 1726  
QY 2619 TGAAGGAGTGAACATCACTTTCATATGCCCCATTTGTTCAACATCCGGAAGACATTTG 2678  
DB 1727 TGAGGAGCATAGTCCCATCTTCTACTCTGCCAGTAGCACCCGCCCCAGAGAAGAACTTG 1786  
QY 2679 TATCACAGCAGAGAGTTGGAGTTGCTG ----- 2708  
DB 1787 TAGCCACAGAGGAGTTGGAGTAGTCTGCTCAAGTAGACCCAGAACAAATCTTTAAT 1846  
QY 2709 ----- 2708  
DB 1847 TGGTTCTTATCTGTGAAGGGACAGGTTAGGGAGTCATAGCTGTCTTTTCTAAAGCC 1906  
QY 2709 -----TTCCCATATACAGATGTCGAGCTTGTGAATGGCAATCCAGTGTGACGG 2759  
DB 1907 CTGCTCTCTCCAGATACACAGAAATTCGCTTGGTGAATGGCAAGACCCCTGTGAGGG 1966  
QY 2760 GCAAGTGAGATCAACGCTTTGGACACTTGGGCTCAGTGTGTGACACCACTGGGACCC 2819  
DB 1967 CAGAGTGAGCTCAAAACGCTTGGTGGCTTGGGGATCCCTCTCTAATCTCACTGGGACAT 2026  
QY 2820 AGAAGATCCCCGTGTTCTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACACAGG 2879  
DB 2027 AGAAGATGCCATGTTCTTTGGCAGCAGCTTAAATGTGGAGTTGCCCTTTCTACCCAGG 2086  
QY 2880 AGGAAATATATTGAGAAAGAAAGTTGCTGTGTGGGACACAGGTTTCATCTGCTTAGG 2939  
DB 2087 AGGAGCAGCTTTTGGAAAGAAATGGTCAGATCTCGAGGCAATATTTTCACTGCACTGG 2146  
QY 2940 GAATGAGTCACTTCTGGATACCTGTCMAATGACACTTCTTGAGCACCTCCCTGTATCCA 2999  
DB 2147 GACTGAGCAGCACATGGGAGATTGCTCTAATCTGCTAGGTGCTTCAATATGCTCTTC 2206

QY 3000 TGGAAATACTGCTCTGTGATCTGCACAGAAAGCCTGACCCAGCAGCTGTTCTTCATGCC 3059  
DB 2207 AGAGCAAGTGGCTCTGTAATCTGCTCAGGAACACAGTCCCAACACTGCTCTCTGTCGA 2266  
QY 3060 CGCAATGTATCTGACCATATTTGCTGCATCTCCAGAGGCGAGTCTTTGATCTGCTT 3119  
DB 2267 TTTCTATCGCTTTTGGGCCCAACAAAGCCTACCATTCCAGAAGAAAGTCTGTGGCCTGCAT 2326  
QY 3120 AGAGGACAAAGCGCTCCGCTAGTAGTGGGACAGCGCTGTCGGGAGAGTAGAGAT 3179  
DB 2327 AGAGAGTGGTCAACTTCGCTTGAATGGAGAGGTGCTGTGCTGGGAGAGTAGAGAT 2386  
QY 3180 CTATCAGCAGCGCTTCTGGGCGACCATCTGTGATGACGGCTGGGACCTGAGGCATGCCCA 3239  
DB 2387 CTATCATGAGGCTCTGGGCGACCATCTGTGATGACAGCTGGGACCTGAGTGTATGCCCA 2446  
QY 3240 CTGTGTGTCTAAAAGCTGGGCTGTGGAGTGGCTTTCAATGCCACGGTCTCTGCTCACTT 3299  
DB 2447 CTGTGTGTGCAGACAGCTGGGCTGTGGAGAGGCCATTAATGCCACTGTTCTGCTCATTT 2506  
QY 3300 TGGGAGGGGTCAGGCGCCATCTGGCTGGATGACCTGAACTGCACAGGAACGAGTCCCA 3359  
DB 2507 TGGGAAGGAACAGGGCCCATCTGGCTGGATGAGATGAATGCAATGGAAGAATCCCG 2566  
QY 3360 TTGTGGCAGTCCCTTCCCGGCTGGGGCAGCAGCACTGCAGGSCAAGAGGAGGACGC 3419  
DB 2567 CATTTGGCAGTCCCATTCACACGGCTGGGGCAGCAAAAATTCAGSCAAGAGGATGC 2626  
QY 3420 AGGGTCTATCTCTCAGAAATTCACAGCTTGGAGCTCTACAGTGAACATGAACAGAGAG 3479  
DB 2627 GGGAGTTATCTGCTCAGAAATTCATGCTGTGAGACTGACCACTGAAGCCAGCAGAGAGGC 2686  
QY 3480 CTGTCTGGGAGATTGGAGTCTTCTATAACGGGACCTGGGCGAGCTCGGAGAGAGAA 3539  
DB 2687 CTGTGAGGGCTCTGGAAGTTTTTACAACTGAGCTTGGGCACTGTTGGCAAGAGTAG 2746  
QY 3540 CATCACACAGCCATGAGGAGCATTTGTGTCAGGAGCTGGGCTGTGGGGAAGATGGAGT 3599  
DB 2747 CATGTCTGAAACCACTGTGGGTGTGTGTGTCAGGAGCTGGGCTGTGCAGACAAAGGAA 2806  
QY 3600 TGTGAGCTCGCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGTGGATGCATATCA 3659  
DB 2807 AATCAACCTTGATCTTTAGACAAGGCCATGCTCAATTCCTATGCGGTGGCAATGTTCA 2866  
QY 3660 GTGCTCTAAAGCAGATATCTCCATATGGCAGTGCCTGTGCCCCATGGAGCGGAGAAAT 3719  
DB 2867 GTGCTCAAAAGACCTGCACACGCTGTGGCAGTGGCCATCATCTCCATGGGCAAGAGACT 2926  
QY 3720 TTTCCAGCCCCAGAGAGACCTGGATCATATGTGAAGATAGAATAAGAGTCCGTGGAGG 3779  
DB 2927 GGGCAGCCCCCTGGAGGAGACCTGGATCATATGTGACAAAGATAAGACTTTCAGGAAGG 2986  
QY 3780 AGACACCGAGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCCTGGGGCACAGTGTG 3839  
DB 2987 ACCCACTTCTCTTCTGGAGCTGTGGAGATCTGGCATGGAGTTCTTGGGGACAGTGTG 3046  
QY 3840 TGATGACTCTGGGACCTGGGCGAGCGGAAGTGTGTGTGACAGCTGGGCTGTGGCTC 3899  
DB 3047 TGATGACTCTGGGACTTGGACGATGCTCAGGTGGTGTGTCAACAACTTGGGCTGTGGTCC 3106  
QY 3900 TGTCTGGCTGCCCTGAGGAGCGCTTGGCTTGGCAGGGAAGTGAACATCTGGTGA 3959  
DB 3107 AGCTTTGAAGCATTTCAAGAAAGCAGAGTTTGGTTCAGGGGACTGGACCATATGGCTCAA 3166  
QY 3960 TGACATGGGTCGCAAGAAATGAGTCAATTTCTATGGGACTCTCAGCCAAACCTGGGG 4019  
DB 3167 TGAAGTGAAGTCAAGGGAATGAGTCTTCTTGTGGGATTTCTCTGCGCAGAGCTGGGG 3226  
QY 4020 ACAGAGTGAAGTGTGACACAAAGAGATGCTGGCTGGAGTGTCTTGGACAGTCTCTGAA 4079  
DB 3227 CCATAGTAGTGTGGGCAACAGAGAGAGCTGCAGTGAATTCACAGATATTTTCAAGTGA 3286  
QY 4080 -----ATCACTGAATGCCTCTCCTCAGGTCATTTAGCACTTTATTTTATC 4121







Db 870 TGGAGTTGCCATCTCCACCCCGGAGGACCACACTTGTGTGAAGAAGGTGATCAGATCCT 929  
QY 1323 GATAAACAGCATATCTTGCACCTGGGAATGAGTCAGCTCTCTGGGACTGCACATATGATGG 1382  
Db 930 RACAGCCGATTTCACTGCTCTGGGCTGAGTCCTCTCTGGAGTTGTCTGTGACTGC 989  
QY 1383 AAAAGCAAAGGAAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTCTCTGATAAGGC 1442  
Db 990 CCTGGGTGCTGACTGTTCCCATGGCAACACAGCCCTCTGTGATCTGCTCAGGAACA 1049  
QY 1443 AGATCTGG----- 1450  
Db 1050 GATCCAGGTGTGTCCTCCAGTGCAAGCACTCCGCTGTCTCAACCTACAGGCTCTCGCGCCCTC 1109  
QY 1451 -----ACCTAAGCCTGTGCGGGCTCATAG 1475  
Db 1110 AGAGGACAGCGCCCTACTGCTCAGACAGCAGGAGCTCCGCTGGTGGAGCGGGCGG 1169  
QY 1476 CCCCTGTTATGGGAGATTTGGAGGTGAATACCAAGGAGAGTGGGGAGCTGTGTGTCATGA 1535  
Db 1170 TCCCTGCGCCGGGAGAGTGGAGATCTTGACCAGGGCTCTGGGACCATCTGTGTATGA 1229  
QY 1536 CAGATGAGCACAAGGAATGAGCTGTTGTGTATAACAATTTGGGATGTGGAAGCCCTAT 1595  
Db 1230 CGGCTGGACCTGGAGATGCCCGCTGGTGTGAGGAGCTGGGCTGTGGAGAAGCCCT 1289  
QY 1596 GCATGTTGTTGATGACCTATTTTAAAGAACATCAGGACCTATTTGGCTGGATGAGCT 1655  
Db 1290 CAATCCACGGGCTGCTCACTTCGGGCAAGGATCAGGGCCCATCTGTTGGACAACCT 1349  
QY 1656 TTCCTGATGGAATGAGTCAAAATATCTGGGACTGTGAAACAGATGGATGGGGAAGCA 1715  
Db 1350 GAACTGCAAGAAAGAGTCCACGCTGTGAGGTGCCCTTCCCGGGCTGGGGCAGCA 1409  
QY 1716 TAATTTGTACACAGAGGATGTGATTTAACCCTGCTCAGTGATGCAACATGGGGCT 1775  
Db 1410 AAATGCAAGACACAGAGGAGCGGGGCTCATCTGCTCAGAGT-----TCCTGGCCCT 1463  
QY 1776 GAGGCTGTGGGGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGATCTTCAAGGAGC 1835  
Db 1464 CAGGATGGTGTGAGGAGCAGCACTGTGCTGGGTGGCTGGAAGTTTCTACAAATGGGAC 1523  
QY 1836 GTGGGCAACAGTGTGATGAGCGGTGGAACAGTAAAGCTGCAGCTGTGGTGTGAGCCA 1895  
Db 1524 CTGGGCAAGTGTCTGCCGTAAACCCATGGAAGACATCACTGTGCACGATCTGCAGACA 1583  
QY 1896 GCTGACTGCCCATCTTATATCATTTGGCATGGGTCTGGAAACGCTTCTACAGGATATGG 1955  
Db 1584 GCTTGGCTGTGGGGACAGTGGAAACCCCTCAACTCTTCTGTGCTTTAGAGAAGGTTTAG 1643  
QY 1956 AAAAATTTGGCTCGATGATGTTTCTGTGATGGAGATGATCAGATCTCTGCTCATGCAG 2015  
Db 1644 GCCACAGTGGGTGATAGATCCAGTTCGGAACACTGACACCTCTCTGCGCATGTC 1703  
QY 2016 GAACAGTGGGTGGGAAATATGATGTCAGTGCAGTCAAGATGTTGGAGTGAATCTGTTTC 2075  
Db 1704 TTCTGACCCCTTGGAAATTAACAATCATGCTCTCAAGAGGAAGCACTATATCTGGT--- 1760  
QY 2076 TGATCATCGGATATGGAGCTGAGGCTTGTGGTGGAGCAGCAGGTGTGCTGGAAAGT 2135  
Db 1761 ---TGACACAGCAGACAGATCCCGCTGTGGATGGAGTGGTCTCTCTGAGAGT 1817  
QY 2136 TGAGTGAATGTCACAGGTGCGGTGGGAATTTCTGTGCTTAATGCTGGGAATCAACAT 2195  
Db 1818 GGAGATCTTGACCAAGGCTCTTGCGGCAACCATCTGTGATGACCCCTGGACCTGGACGA 1877  
QY 2196 TGCTGAAGTTGTTGACGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGACC 2255  
Db 1878 TGCCCGTGTGTGTCAGCAGCTGGGCTGTGGAGAAGCCCTGGAGCCCATGCTCTCTTC 1937  
QY 2256 TCATTTACAGAAAGAAATACATCTTAATGTCGAATTTCTGGCTGCATGGAGGGA 2315  
Db 1938 CTTCTTGGGAGGGATCAGGGCCCATCTGGCTGGATGAAAGTGAACCTGCAGAGGAGGA 1997

QY 2316 AGCCTCTCTCTGGGATTTGTATACGATGGGAGTGGAAACACAGACTGGTGTGTCATTTAAATAT 2375  
Db 1998 CTCCCAAGTATGGAGTGCCTTCTCTGGGATGGCGCAACAACATGCAATCATCAAGA 2057  
QY 2376 GGAACAAGTTTGTATCTGCTCAGCCCCACAGGAGCCAGGCTGGTGGAGCTGATATGCC 2435  
Db 2058 AGATCAGAGGATCATCTGCTCAGGATTTGTGC-----GTCCTGGCTGGAGAGATGGACC 2111  
QY 2436 CTGCTCTGACGCTGTTGAAGTGAACATGCACACATGCGCTCTGTCTGTGATTTCTGA 2495  
Db 2112 CTGCTCAGGGCGAGTAGAAGTGCATTTGGAGAAGCTGGACCCAGTCTCTGATGGAAA 2171  
QY 2496 TTTCTCTCTTCACTGCTGCAATGCTGTGCAGAGAAATTAATTTGGAGATGCCATATC 2555  
Db 2172 CTTTCACTCTCCCACTGCCAGGCTATCTGTGCAGAGCTGGGATGTGGCAAGCTGTGTC 2231  
QY 2556 TCTTTCTGTGGAGATCACTTTGGAAAGGGAATGCTTAACCTTGGGCGGAAAAGTTCCA 2615  
Db 2232 TGTCTGGGACACATGCCAATTCAGAGATGCCATGGGCTGGAAGAGTTTCAAG 2291  
QY 2616 GTGTGAAGGGAGTGAACCTCACTTGCATTAATGCCCCATTTGTTCAACATCCCGAAGACAC 2675  
Db 2292 GTGTGATGGGGGAGCTGAGCTCTGTCTGCCAGAGTGCCTGTCTCCAGGAGGAC 2351  
QY 2676 TGTATCCACACAGCAGAGATTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGT 2735  
Db 2352 ATGTCTCCACAGTGGAGTGTCTCAGTTGTCTGTTCAAGTTACACAGAAGTCCAGCTTAT 2411  
QY 2736 G---AATGGCAATCCCACTGACGGGCAAGTGAGATCAACGCTGTGGACACTGGGG 2792  
Db 2412 GAAAAACGCGCTCTCAATGTGAGGGCAGGTGGAGATGAAGATCTCTGGACGATGGAG 2471  
QY 2793 CTCACCTGTGTACACCCCACTGCGGACCCAGCAAGATGCCGCTTCTATGACAGACAGTCA 2852  
Db 2472 AGCGCTGTGTGCTCCCACTGGAGTGTGCCAATGCCAATGTGTCTGTCTGCTGAGTCCGG 2531  
QY 2853 CTGTGGGACTGTCTCTCAACACACAGGAGGAAAATATATTGGGAAGAAGTCTTCTGTGT 2912  
Db 2532 CTGTGGAGTGCCTCTCCACCCCGAGAGGACCACACTTGGTGGAGGAGGTGATCAGAT 2591  
QY 2913 GTGGGACACAGGTTTCAATTTAGGGAATGAGTCACTTCTGGATAACTGTCAATATGAC 2972  
Db 2592 CTCAACAGCCCAATTTCACTGCTCAGGGCTGAGTCTCTCTGTTGGAGTTGCTCTGAC 2651  
QY 2973 AGTTCTGTGAGCAGCTCCCTGTATCCATGGAATACTGTCTGTGATCTGCACAGAGAG 3032  
Db 2652 TCCCTTGGTGGGCTGACTGTTCCATGGCAACACAGGCTCTGTGATCTGTCTAGGAAA 2711  
QY 3033 CTGACCCAGCCACTGTTTCCATGCTCGCAATGTATCTGACCCCATATTTGTCTGACGT 3092  
Db 2712 CCACACCCAGGTGCTGCCCGAGTGCACACACTTCTGTCTCAACCTGACGGCTCTCGGC 2771  
QY 3093 TCACAGGCGACTGTTTGTATCTGTAGAGCAACAGGCTCCGCTAGTGGAGGGA 3152  
Db 2772 CTCAGAGGAGATTTCTCCCTACTGCTCAGACAGCAGGAGCTCCGCTGTGGCGGGG 2831  
QY 3153 CAGCCGCTGTGCGGAGAGTAGATCTATCAGCAGGCTTCTGGGGCACCATCTGTGA 3212  
Db 2832 CGTCCCTCGCGGGGAGAGTGGAGATCTTGACCAGGCTCTGGGGCACCATCTGTGA 2891  
QY 3213 TCACGCGCTGGGACCTGAGCGATGCCACAGTGTGTGTCTCAAAAAGCTGGGTGTGGAGTGGC 3272  
Db 2892 TGATGACTGGGACCTGGAGATGCCGCTGTGTGTGTCAGGCAAGCTGGGTGTGGAGAGC 2951  
QY 3273 CTTCAATGCCAGGCTCTGTCTACTTTGGGAGGGGTGAGGCCCATCTGGCTGATGA 3332  
Db 2952 CTTCAATGCCAGGCTCTGTCTACTTCCGGGCAAGGATCAGGGCCCATCTGGCTGAGCA 3011  
QY 3333 CTTGAATGCACAGAAAGGAGTCCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGGCA 3392  
Db 3012 CTTGAATGCACAGAAAGGAGTCCCACTGTGTGAGGTGCCCTTCCCGGGCTGGGGGCG 3071

Query Match	14.0%	Score	649.8	DB	4	Length	2305
Best Local Similarity	64.7%	Pred. No.	4.7e-151				
Matches	1039	Conservative	0	Mismatches	45	Indels	24
Gaps	4						
Qy	2473	TGGCGCTCTGCTGTGATCTGTGATTTCTCTTCATCGTCCCAATGTGCTCTGCAGAGAA	2532				
Db							
Qy	4	TGGACCCAGTGTCTGATGGGAATTTACATTCCTCCCACTGCACAGGTCATCTGTGCAGAA	63				
Db	64	CTGGGATGTGGCAATGTGTGTCTGTCTCTGGACACGTGCCCTTCACAGAGTCCAGATGCG	123				
Qy	2593	CTAATCTTGGGCGGAAAAGTTCACGTGTGAAGGAGTCAAACTCACCTTGCATTATGCCCC	2652				
Db							
Qy	124	CTGTCTCTGGGCTGAAGAGTTTCAGGTGTAAAGGAGAGAGGCTTGAGCTCTGGCTGCGCCC	183				
Db							
Qy	2653	ATTCTTCAACATCCGGAAGACACATTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTCC	2712				
Db	184	AGATGGCCCTGTCCAGGGGCGCATTTGTCAACCACAGTGGACTGTTTCATGTTGTCGTT--	241				
Qy	2713	CGATATACAGATGCCGACTTTGTG---AATGGCAATCCCGAGTGTGACGGGCAAGTGGAG	2769				
Db	242	-----CAGAAATGCCACTCATGAAAAATGGCTCTCTCAATGTGAGGGCGAGGTGGAG	294				
Qy	2770	ATCAACCTGCTTGACACTCGGGCTCACTGTGTACACCCACCTGGGACCCAGAGATGCC	2829				
Db	295	ATGAATATTTCTGGACCGTGGAGAGCACTCTGTGCCCTCCCACTGGAGCTGTGCCCAATGCC	354				
Qy	2830	CGTGTCTATGCACAGCTCAGCTGTGGGACTGCTCTCTCAACACAGGAGGAAAAATAT	2889				
Db	355	AATGTTCTGTCTCAGCTCGGCTGTGGAGTGCCTATCTCCACCCCTAGAGGACCACAC	414				
Qy	2890	ATTGGAGAAAGATGTTGTTGTGTGGGGACACAGTTTTTCATTTGCTTAGGAATGAGTCA	2949				
Db	415	TTTGTGGAAGGAGTGATCAGATCTCAACAGTCCGATTTTCTGTTTCCGGGGTGGAGTCC	474				
Qy	2950	CTTCTGTGATAACTGTCAAAATGACAGTTTCTTGGAGCACCTCCCTGTATTCATGGAATACT	3009				
Db	475	TTTCTGTGGAGTTGTCCTGTGTGACTGTCTGGTGGTCTCTGACTGTTCCTCCATGGCAATG	534				
Qy	3010	GTCTCTGTATCTGCACAGGAAGCCTGACCCAGCCACTGTTTCCATGCCTCGCAAAATGTA	3059				
Db	535	GCCTCTGTATCTGCTCAGGAACACAGCCAGGTGTGCCCGACGTCAACAGCTCCATG	594				

Db	242	-----CAGAAGTCCGACATCATGAAAAATTGGCTCCCTCAATGTGTAGGGGCGAGGTGGAG	294
Qy	2770	ATCAACCTGCTTGACACTGGGGTCTACTGTGTGACACCCACTGGGACCCAGAGAAGATGCC	2829
Db	295	ATGAAATATTTCTGGACGGTGGAGAGCACTCTGTGCCTCCCACTGGAGTCTGTGGCCAATGCC	354
Qy	2830	CGTGTCTTATGCAGACAGCTCAGCTGTGGGACTGCTCTCTCAACACCAAGGAGGAAAAATAT	2889
Db	355	AATGTTCTCTGCTCAGCTCGGCTGTGGAGTGCACCATCTCCACCCCTAGAGGACCAACAC	414
Qy	2890	ATTGGACAAGAAGTGTTCGTGTGTGGGACACAGGTTTCATGCTTTAGGGAATGACTCA	2949
Db	415	TTTGTGGAAGAAGTGAATCAGATCTCAACAGTCCGATTTTCACTGTTCCGGGGTGGAGTCC	474
Qy	2950	CTTCTGGATAACTCTAAATGACAGTTCTTTGGAGCACCTCCCTGTATTCATGGAAATACT	3009
Db	475	TTTCTGTGGAGTTGTCTCTGTGACTGCTCTGGTGGTCTTGACTGTTCCTCCATGGCAACATG	534
Qy	3010	GTCCTCTGTATCTGCACAGGAAGCCTGACCCAGGCATCTGTTTCCATCGCTTCGCAAAATGTA	3059
Db	535	GCCTCTGTATCTGCTCAGGAACACAGCCAGGTGTGCCCGACGTCAACAGCTCCATG	594

QY 3070 TCTGACCCATATTTCTCTCAGTTTCCAGAGGSCAGTCTTTGATCTGCTTAGAGACAAA 3129  
|||||  
Db 595 TCTGAAGAGCAGGCTCTGTAGTCTCAGAAGAGATCGCCCCCTACTGCTCAGATAGCAGA 654  
QY 3130 CGGCTCCGCCCTAGTGGAGTGGGACAGCGCTGTGCGGGAGAGTAGATCATCACGAC 3189  
|||||  
Db 655 CAGCTCCGCCCTGGTGGGAGGAGTCCCTGTGCGGGAGAGTAGATCATCACGAC 714  
QY 3190 GGTCTCTGGGACCATCTGTGATCAGCGCTGGGACCTGAGGATGCCACGTGTGTGT 3249  
|||||  
Db 715 GGTCTCTGGGACCATCTGTGATCAGCGCTGGGACCTGAGGATGCCACGTGTGTGT 774  
QY 3250 CAAAAGCTGGGCTGTGGAGTGGCTTCAATGCCAGGCTCTCTGCTCACTTTTGGGAGGG 3309  
|||||  
Db 775 AGGCAGTTGGGCTGTGGAGAGCCCTCAATGCCAGGCTCTCTGCTCACTTGGGAGG 834  
QY 3310 TCAGGGCCCATCTGCTGGATGACCTGAACCTGCACAGAGAGGATGCCACTTGTGGAG 3369  
|||||  
Db 835 TCAGGGCCCATCTGCTGGATGACCTGAACCTGCACAGAGAGGATGCCACTTGTGGAG 894  
QY 3370 TGCCCTTCCCGGGCTGGGAGCAGCAGCTGCAGGACACAGGAGGAGCGGGCTCATC 3429  
|||||  
Db 895 TGCCCTTCCCGGGCTGGGAGCAGCAGCTGCAGGACACAGGAGGAGCGGGCTCATC 954  
QY 3430 TGCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTCTGG 3489  
|||||  
Db 955 TGCTCAGAAATTCACAGCCTTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTCTGG 1008  
QY 3490 AGATTGGAGTCTTCTTAACGGGACCTGGGACAGCTGGGAGGAGGAGGAGGAGGAGG 3549  
|||||  
Db 1009 TGGCTGGAAGTTTCTTACAAATGGGACCTGGGAGCAGTCTGCCGAGGAGGAGGAGG 1068  
QY 3550 GCATAGCAGGAGTGTGTGAGGAGCAGTGGGCTGTGGGAGAGAGTGTGAGGAGGAG 3609  
|||||  
Db 1069 ATCACTGTGCTCATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1128  
QY 3610 GCCCTTTTATACAGAGGCTGTGTTTATGTGGGTGGATGACATTCAGTGTCTCTAAA 3669  
|||||  
Db 1129 TCTGTGTCTTAGAGAGGTTTATAGACCTCGTGGGTAGATGGAATCCGGTGTGCGAAG 1188  
QY 3670 AGCATATCTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3729  
|||||  
Db 1189 ACAGATATCTCTCTGAGCAGTGTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1248  
QY 3730 GCAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3783  
|||||  
Db 1249 AAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308  
QY 3784 ACCGAGTCTCTGGAGAGTGGAGATCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 3843  
|||||  
Db 1309 GGTCTCTGGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 1368  
QY 3844 GACTCTGGGAGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 3903  
|||||  
Db 1369 GACAGTGGGAGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 1428  
QY 3904 CTGGCTGGGAGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 3963  
|||||  
Db 1429 CTCAGTGGGAGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 1488  
QY 3964 ATGGGCTGCAAGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 4023  
|||||  
Db 1489 CTGAAGTGGGAGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 1548  
QY 4024 AGTACTGTGACAGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 4068  
|||||  
Db 1549 CACGACTGTGGAGGAGGAGTGGAGATCTCTGGACAGGCTCTCTGGGAGCAGATGTGTGAT 1593

RESULT 12

SSSRP3  
LOCUS

SSSRP3 1665 bp DNA linear MAM 27-AUG-1997

DEFINITION S.scrofa mRNA for scavenger-receptor protein, 1665bp.  
ACCESSION X99334  
VERSION X99334.1 GI:1480360  
KEYWORDS WC1 gene.  
SOURCE Sus scrofa.  
ORGANISM Sus scrofa  
REFERENCE 1 (bases 1 to 1665)  
AUTHORS Kanan, J.H., Nayeem, N., Binns, R.M. and Chain, B.M.  
TITLE Mechanisms for variability in a member of the scavenger-receptor  
cysteine-rich superfamily  
JOURNAL Immunogenetics 46 (4), 276-282 (1997)  
MEDLINE 97364683  
PUBMED 9218529  
REFERENCE 2 (bases 1 to 1665)  
AUTHORS Kanan, J.H.C.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-1996) J.H.C. Kanan, University College London,  
Immunology, Fourth floor, Windeyer Building, Cleveland Street,  
London WC1 6BT, UK  
FEATURES  
source  
1..1665  
/organism="Sus scrofa"  
/strain="SLAB/b"  
/db\_xref="taxon:9823"  
/cell\_type="peripheral blood T lymphocyte"  
29..1665  
/gene="WC1"  
29..1665  
/gene="WC1"  
/codon\_start=1  
/product="scavenger-receptor protein"  
/protein\_id="CAA67708.1"  
/db\_xref="GI:1480361"  
/db\_xref="SPTREMBL:Q29111"  
/translation="MALDRHLSLQGLFLLIVVGQALRLKDRHRCGRVELKH  
QKMGTVNAYWGLPFAVVCQLECGTAVDAPVASYFGRGDGPILWAFEGESTL  
TDCRPNITDYRTKGFSDMDAGAVCSGFLVVGDRPCSGRVEVHSGEDMTPVSDGN  
FTPTAHVCAELGCGKAVSVLQGPRLPGQVNMWPEERCKDEPELWSCPVPVPCPV  
GTCFHSAGAVOVCSNAYSEVRLMRNGTSOCAGRVEMN1SGGHRTLCASHWTANAVVC  
HQLCGVALSTPKGNEGDDQMKARPHCSGAEFLWNCPTVALGVPDCAHGNTASVIC  
SGNTQVLQPCNDLSLEPAGSAASESTNCSESRLRLVDESYSRRVELLHOGSW  
TICDESDLDLDAHVCRQLCGDAISALGAFHGVGSGRIWLDVNGCTGKESHLWQC  
PFRGWRDHRKEDAGVICSLEALRMVSEDCAGWLEVFYNGTWGVCGRSPMEDI  
TLSMICRQLGCGDSTLNSVSLRSGSRPRWVDGIQC"  
29..94  
sig\_peptide  
/gene="WC1"  
BASE COUNT 344 a 405 c 539 g 377 t  
ORIGIN  
Query Match 11.9%; Score 551.8; DB 4; Length 1665;  
Best Local Similarity 61.6%; Pred. No. 1.4e-126;  
Matches 974; Conservative 0; Mismatches 587; Indels 20; Gaps 5;  
QY 2086 GATATGGAGCTGAGGCTGTGGTGGAGCAGCAGCTGTGCTGGAAGAGTTGAGGTGAAT 2145  
|||  
Db 98 GCTCTGGAGCTGAGCTGAAGGATGGAAGCACCCTGTGAGGGAGAGCTGGAATTTGAAG 157  
QY 2146 GTCCAGGGTCCGCTGGGAATTTCTGTGCTAATGCTGGGGAATGAACATTTGCTGAAGTT 2205  
|||  
Db 158 CACCAAGGAAAATGGGGCAGCTGTGAATCAGGCTTACTGGGGTTTCCAGGAGCAACCGTG 217  
QY 2206 GTTTCAGGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCCAGAGAGCTCATTTTCA 2265  
|||  
Db 218 GTGTGCACACAGCTGGAATGTGAACTGCTGTGATGCTCCCGTGGCATCTTTATTTGA 277  
QY 2266 GAAAGAACATTACATCTTAATGTGCAATTTCTGCTGCACTGGAG-GGGAAGCCCTCT 2324  
|||  
Db 278 CGAGGGGATGGTCCCATCTGGTTGTATGCTTTTGTGAAGGACAGAGTCAACTCTCACT 337  
QY 2325 CTGGGATTTATACATGGGAGTGGAAACAGACTCGGTGCTCATTTTAAATATGGAAGCAAG 2384  
|||

Db 338 GACTGTAGGCATCTTAATATACAGACTACAGGACTAAAGGCTTTTCCCATGATTGGAT 397  
QY 2385 TTGTGATCTGCTCAGCCACAGCCAGCCAGCTGTTGAGCTGATATGCTGCTCTGG 2444  
Db 398 GCTGGAGAGTCTGCTCAGGATTTGTGC--GCCTGGTTGAGGGAGACAGACCTTGTCTCAGG 456  
QY 2445 ACCTGTTGAAGTGAACATATGACACACATGCGCGCTGCTGCTGATGATTTCTCTCTCT 2504  
Db 457 GCGAGTGAAGTGCATCTGAGAGAGACTGGAGCCAGTGTCTGATGGGAACCTTACAT 516  
QY 2505 TCATGCTGCAATGCTCTGTCAGAGAAATTAATGTCGAGATGCCATATCTCTCTCTG 2564  
Db 517 CCCCAGTCCCACTGCTGTCAGAGCTGGGATGCGCAAGGCTGTGCTGCTCTGG 576  
QY 2565 GGGAGATCACTTTGAAAAGGGAATGCTTAACCTGGCGCGAAAGTCCAGTGTGAAGG 2624  
Db 577 TCAGAGCCCTCAGAGAGCCAGGTAACCAAGTCTGGCTGAAGAGTTCAGGTGTAAAGG 636  
QY 2625 GAGTGAATCTCACTTGCATTTATGCCCCATTTGTTCAACATCCGGAAGACACTTGTATCCA 2684  
Db 637 GGAGAGCCCTGAGCTCTGCTGCTGCCAGAGTGCCTGTCCAGGAGGCACTTGTCAACCA 696  
QY 2685 CACAGAGAGTGGAGTGTCTGTTCCCGATATACAGATGCCAGTCTGTCGCTCTG 2741  
Db 697 CAGTGGAGCTGTTCAAGTGTCTGTTTCAGCATACTCAGAAAGTCCGGCTCATGAGAAATGG 756  
QY 2742 CAAATCCAGTGTGACGGGCAAGTGCAGATCAACGTGCTTGGACACTGGGGCTCACTGTG 2801  
Db 757 CACTCTCAGTGTGCGGGGAGTGGAGATGATATTTCTGGAGGTGGAGACACTCTG 816  
QY 2802 TGACACCCACTGGGACCCAGAAAGATGCCGCTGTTATGTCAGACAGACTCAGCTGTGGGAC 2861  
Db 817 TGCTCCCACTGACCATGACCAATGCCAATGTTGTGTCATCAGCTGGCTGTGGAGT 876  
QY 2862 TGCTCTCAACACAGGAGGAAATATATTGAGAAAGAGTTCGTTGTTGGGGACA 2921  
Db 877 CGCCCTCTCCACCCGGAAGGA-----GCAGAGGAGGTGATCAGATGTGGAAGC 927  
QY 2922 CAGTTTCATTTGTTAGGGAATGAGTCACTTCTCGGATAACTGTCAATGACACTTCTTGG 2981  
Db 928 CCGATTTCACTGCTCTGGGCTGAGTCTTCTTGTGGAATTTGCTGACTGCCCTGGG 987  
QY 2982 AGCACTTCCCTGTATCCATGAATATCTGCTGTGATCTGCACAGAGAGCCCTGACCCA 3041  
Db 988 TGTTCCTGACTGTGCCATGAAACACAGACCTCTGTGATCTGCTCAGGAACACAGACCCA 1047  
QY 3042 CCACTGTTTCCATGCTCCCAATGATCTGACCCATATTTCTGCACTTCCAGAGGG 3101  
Db 1048 GGTGCTGCCTCAGTCAATGACTCCCTGTCTGAACACAGAGGCTCTGCACCCCTCAGAGGA 1107  
QY 3102 CAGTCTTTGATCTGTAGAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTG 3161  
Db 1108 AAGCACCACCACTGCTCAGAGAGCAGACAGCTGCGCTGTTGATGAGAAAGTATTG 1167  
QY 3162 TGGCGGAGAGTAGAGATCTATCAGACAGCGCTCTTGGGGCACCATCTGTGATGACGCTG 3221  
Db 1168 CTCACGGAGAGTGGAGATCTTCAACCAAGGCTCTGCGGGCACCATCTGTGATGAAAGCTG 1227  
QY 3222 GGACTGAGCGATGCCACCTGCTGCTCAAAAGCTGGCTGTGGAGTGGCTTCAATGC 3281  
Db 1228 GGACTGGAGATGCGCACGTGGTGTGTCAGGCACTGGCTGTGGAGATGCCATCAGTGC 1287  
QY 3282 CACGGTCTCTGCTCAGTTTGGGAGGGGTGAGGGCCCATCTGCGTGGATGACCTGAAGTGC 3341  
Db 1288 ACTGGGTCTGCTCACTTTGGGTGGGCTCAGAGAGCATCTGGCTGATGACCTGAAGTGC 1347  
QY 3342 CACAGGAACGAGTCCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGGGCAGCAGACTG 3401  
Db 1348 TACAGGAAAGAGTCCCACTATGGCAGTGCCTTTTCAGGGGCTGGGGGGCAGCAGACTG 1407  
QY 3402 CAGCACAAGGAGGACGCGGCTCATCTGCTCAGATTCACAGCCTTGTGAGCTCTACAG 3461  
Db 1408 CAGACACAAGGAGGACGCGGGGTCTCTGCTCAGATCTGCTGAGTCCCTGGCCCTCAGG-----AT 1461

TITLE  
JOURNAL

QY 3462 TGAACHTGAACAGAGAGCTGTGCTGGAGATTTGGAAGTCTTCTATACGGGACCTGGGG 3521  
Db 1462 GGTGAGCGAGGACAGAGTGTGCGGGTGTGCTGAAGTTTCTTACAGGGACCTGGGG 1521  
QY 3522 CAGCTCGCGGAGGAGAAATCACCACAGCCATACAGCATTGTGTCAGGACCTGGG 3581  
Db 1522 CGCTGTGCGCCACCCCATCACTTGTCCATGATCTCGAGCAGCTGG 1581  
QY 3582 CTGTGGGAGATGAGTTGTCAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTTCAT 3641  
Db 1582 CTGTGGGACATGGAACCTCACTTCCGTTTCTTCTAGGAGGTCTAGACCCG 1641  
QY 3642 GTGGGTGATGATTCAGTG 3662  
Db 1642 GTGGGTGATGATTCAGTG 1662

RESULT 13  
AC131207/cLOCUS  
DEFINITION  
AC131207  
Homo sapiens clone RP11-326A16, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 10  
unordered pieces.ACCESSION  
AC131207.1 GI:22296912  
VERSION  
HTG: HTGS\_PHASE1.KEYWORDS  
SOURCE  
human.ORGANISM  
Homo sapiensEukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.REFERENCE  
AUTHORS

1 (bases 1 to 202252)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, I.,  
Barbieri, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,  
Boulay, J., Bowles, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
Burch, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Dunthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Galsi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlssohn, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisegh, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, B., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Oragunye, N., Nickerson, E., Nwokoko, S., Ogih, M., Okuwonu, G.,  
Oquendo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivers, M., Rojds, A., Rojokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G., and Gibbs, R.

Direct Submission  
Unpublished



```
/codon_start=1
/product="EP4a"
/protein_id="AAK92475.1"
/db_xref="GI:15217047"
/translation="RRSLHCPISLSRQIRLVDDGGSCVGRVEILDDQSGWGTICD
DWDLDARVCRQLGCEALNATGSAHFAGSGGPILWLDLNCAGNESHVWPCRSRGW
GRDCRHRKEDAGVICSEYLALRMVSEDOECAGLVEFYNGTWGSCVCRSPMDVTYSII
CSQLGCGSGSLNTSVGLREGSRPHIDLQCRKTDLSLWQCPSPWFFSSCPKEEA
YISAGRRPKQAPAAAPCTDRKRLRGDSECSGRVEVHWSGWSLWDDSDSLAEA
EYVCOQGGQALAVRAGTDFPGNSIWLDEVRGREGSSLWDCAAEPWGGSDCKHE
EDAGVRCYSVRKTLPTTPAGTINIGHLERILRSGLDFGCSLTAQERKPEEPWLVGRS
PGRYVELLCLGWSGHTFLMSVQCRI"
BASE COUNT      420 a 566 c 707 g 470 t      2 others
ORIGIN
Query Match      10.3%; Score 476.4; DB 4; Length 2165;
Best Local Similarity 69.5%; Pred. No. 1e-107;
Matches 701; Conservative 1; Mismatches 252; Indels 54; Gaps 2;
QY 3108 TTTGATCTGTTAGAGACAAAGCGTCCGCCCTAGTGGATGGGACAGCGCTGTGCCGG 3167
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 TTTCTCTCTCTACAGACAGACAGATCCGCTGTGGAGCGGGCGGTTCCTGCGTCGG 86
QY 3168 GAGATGAGATCTATACAGACGCTTCTGGGACCATCTGTGATGACGCTGGACCT 3227
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
87 GAGATGGAGATCTTGACAGGCTCTCTGGGACCATCTGTGATGACGCTGGACCT 146
QY 3228 GAGCGATGCCACGCTGTGTCAAAAGCTGGCTGTGGATGGCTTCAATGCCACGCT 3287
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
147 GGAGATGCCCGCGTGTGTGACGAGCTGGCTGTGGAGAGCCCTCAATGCCACGG 206
QY 3288 CTCTGCTCACTTGGGAGGGGTGAGGCCATCTGGCTGGATGACCTGAATGACAGG 3347
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 GTCTGCTCACTTGGGAGGATCAGGGCCATCTGGCTGGAGACCTGAATGCGCGCG 266
QY 3348 AACGGAGTCCACTTGTGGAGTGCCTTCCCGCGCTGGGGCAGCAGCTGCAGGCA 3407
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267 AATGATCTCAGTGTGGAGTGCCTTCCCGGGCTGGGGCGGACGACCTGCAGACA 336
QY 3408 CAAGGAGGAGCGGGGTCTATCTGCTCAGAAATTCACAGCTTGAGCTCTACAGTAAAC 3467
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
327 CAAGGAGGAGCGGGGTCTATCTGCTCAGAGTACCTGGCCCTCAGG-----ATGCTGAG 380
QY 3468 TGAACAGAGAGCTGTCTGGAGATGGAAGTCTTCTATAAGCGGACCTGGGGCAGCGT 3527
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
381 CGAGGAGGAGTGGCTGGGTGGTGAAGTTTCTTCAACAGGACCTGGGGCAGTGT 440
QY 3528 CGGAGGAGGAGCATCACACAGCCATAGCAGGATTTGTGACGAGCTGGGCTGTGG 3587
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
441 CTCGCGAGCGCCATGATGATGATCACCCTGTCCATCATCTGCAGTCAGCTTGGCTGTGG 500
QY 3588 GGAGATGGAGTGTGACGCTCGCCCTTTATCTAAGACAGGCTCTGTTTCATGTGGGT 3647
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
501 GGACAGTGAAGTCTCAACACTTCTGTGCTCTCAGGAGAGTCTTAGACCCCGGTGGAT 560
QY 3648 GGATGACATTCAGTCTTAAAGCATATCTCCATATGCACTGCTGTCTGCCCATG 3707
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
561 AGATTAATTCAGTGTGGAAGAACTGATACCTCTCTGTCGATGCTCTTCTGGCCCGTG 620
QY 3708 GGAGCAGAAGATCTCCAGCCACAGAGAGACCTGGATCAGATGTGAAGATAGAATAA- 3766
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
621 GAAATTCAGTCTCTCTCCAAAGGAGGAGGCTACATCTCTCTGTGTCAGGAAGAAGACC 680
QY 3767 -----GAGTCGCTGGAG 3779
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
681 CAAGAGCTGTCCAGCTGCTGCCCTCTGCACAGACAGAGAGAGCTCCGGCTCAGGGGAGG 740
QY 3780 AGACACCGAGTGTCTGGGAGAGTGTGAGATCTGGCAGCAGGCTCTCTGGGACACAGTGTG 3839
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
741 AGACACCGAGTGTCTAGGCGGGTGTGAGTGTCTGGCAGACGGCTCTCTGGGACACCGTGTG 800
QY 3840 TGATGACTCTCTGGGACCTGGCGGAGCGGAAGTGTGTCTCAGCAGCTGGGCTGTGGCTC 3899
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 801 CGATGACTCTCGAGCCTGGCAGAGCGCGAGGTGGTGTGTCTCAGCAGCTGGCTGTGCCA 860
QY 3900 TGTCTTGGTGCCTCAGGAGCGCTTCGTTTGGCCAGGGAACCTGGAACCATCTGGTTGGA 3959.
Db 861 GGCCCTGGAAGCGCTCGCGCTGGGACATTCGCCCTTGGAAATGGGAGCATCTGGCTGGA 920
QY 3960 TGACATCGCGTCAAGAGAAATGAGTCATTTCTATGGGACTGTACGCCCAACCCCTGGG 4019
Db 921 CGAGGTGCGGTGAGGGGGGGGAGTCTCTCCCTGTGGGACTGTGCTGCCAGCCCTGGG 980
QY 4020 ACAGATGACTGTGGACACAGGAAGATGTCGCGCTGAGGTGCTGCTGG 4067
Db 981 GCAGACGACTGCAARACAGGAGGACGCTGCTGTGAGGTGCTCTGG 1028

RESULT 15
SSSRP2      1316 bp      DNA      linear      MAM 27-AUG-1997
LOCUS      S.scrofa mRNA for scavenger-receptor protein, 1316bp.
DEFINITION X99333
ACCESSION X99333.1 GI:1480358
VERSION   WCI gene.
KEYWORDS  Sus scrofa.
SOURCE    Sus scrofa.
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE  1 (bases 1 to 1316)
AUTHORS   Kanan,J.H., Nayeem,N., Binns,R.M. and Chain,B.M.
TITLE     Mechanisms for variability in a member of the scavenger-receptor
JOURNAL   Cysteine-rich superfamily
MEDLINE   Immunogenetics 46 (4), 276-282 (1997)
PUBMED    9218529
REFERENCE  2 (bases 1 to 1316)
AUTHORS   Kanan,J.H.C.
TITLE     Direct Submission
JOURNAL   Submitted (15-JUL-1996) J.H.C.-Kanan, University College London,
Immunology, Fourth floor, Windeyer Building, Cleveland Street,
London WC1 6BT, UK
FEATURES
source
1..1316
Location/Qualifiers
/organism="Sus scrofa"
/strain="SLAb/b"
/db_xref="taxon:9823"
/cell_type="peripheral blood T lymphocyte"
<1..126
/number=1
join(127..219,301..1316)
/number="WCI"
join(<127..219,301..>1316)
/number="WCI"
/codon_start=3
/product="scavenger-receptor protein"
/protein_id="CAA67707.1"
/db_xref="GI:1480359"
/db_xref="SPTREMBL:Q29110"
/translation="NQTVLPQCNDSLSEPAASAEECTANCSDSRLRLVDGSSYC
SGSVETLHOGSMGTVCDDSDLDHVCRLQCGKTLRSALGSAHFAGSGRIWLDLV
NGTGVSRHLMQCPSPRGWQHNRHKEDAGVICSEYLALRMVSEDOECAGLVEFYNGT
WGVGRCRSPMEDITLITICRLQCGSGTINSSVLRGSRPWDRICQKRTDSSLWE
CPSPDKYSSCSKEEAYVSCAGRRPKSCPTTAPCTDKELRLRGDSCVSGRVEVWH
GGSWCTVCDSDSLAEAEVVCQQLGCGHALEALRDAAAFPGSGSIWLDEVQCRGRESS
LWDCAAGPWGQSDCKHEADAGVRCGARGPSNP"
127..219
/number="WCI"
/number=2
220..300
/number="WCI"
/number=2
301..>1316
/number="WCI"
/number=3
BASE COUNT      264 a 345 c 422 g 285 t
```

## ORIGIN

Query Match 10.0%; Score 460.6; DB 4; Length 1316;  
Best Local Similarity 68.8%; Pred. No. 8.7e-104;  
Matches 690; Conservative 0; Mismatches 259; Indels 54; Gaps 2;

Search completed: May 12, 2003, 05:29:51  
Job time : 8566 secs

QY 3113 TCTGCTTAGAGACAAACGGCTCGCCTAGTAGTGGAGACAGCCGTGTGCCGGGAGAG 3172  
DB 292 TCTATTACAGACAGACAGACAGTGGCCCTGTGTGGATGGAGGAGTTACTGTCTCAGGGAGCG 351  
QY 3173 TAGAGATCTATCAGACGGCTTCTGGGGCCACCCTCTGTATGACGGCTGGGACCTGAGCG 3232  
DB 352 TGGAGATCCTTACAGACGGCTTCTGGGGCCACCCTGTGTATGACAGTGGGACCTGGAGC 411  
QY 3233 ATGCCACAGTGGTGTCAAAAGCTGGGCTGTGGAGTGGGCTTCAATGCCACGGTCTCTG 3292  
DB 412 ATGGCGCATGTGGTGTGAGGACAGTGGCTGTGGAANAACCTCAGTGCACCTGGGCTCTG 471  
QY 3293 CTCACCTTTGGGAGGGGTACAGGGCCCATCTGGGCTGGATGACCTGAATGCACAGGAACGG 3352  
DB 472 CTCACCTTTGGGAGGGGTACAGGGCCCATCTGGGCTGGATGACCTGAATGTACAGGAAGG 531  
QY 3353 AGTCCCACTTGTGCACTGGCTTCCCGCGCTGGGGCCAGCAGCTGCAGGCACAAGG 3412  
DB 532 AGTCCCACTTATGGCAGTGGCTTCCCGGGCTGGGGCCAGCACAATGCAGACACAAGG 591  
QY 3413 AGGACGAGGGGTCTATCTGCTCAGAAATTCACAGCTTTCAGGCTCTACAGTGAACCTGAAA 3472  
DB 592 AGGACGAGGGGTCTATCTGCTCAGATCCCTTGGCCCTCAGG-----ATGGTACGCGAGG 645  
QY 3473 CAGAGAGCTGTGTGGAGATTTGAAGTCTTCTATACGGGACCTGGGGCAGCCTCGGCA 3532  
DB 646 ACCAGAGTGTGGCGGGTGGCTGGAGTCTTCTACACGGGACCTGGGGCGGTCTCTGCC 705  
QY 3533 GGAGGAACATCACACACAGCCATGACAGCATTTGTSCAGGCAGCTGGGCTGTGGGGAGA 3592  
DB 706 GCACCCCATGGAGAGATCACCTTGTCCATCATCTCGAGGACGCTTGGCTGTGGGGACA 765  
QY 3593 ATGGAGTGTGACGCTCGCCCTTTATCTAAGACAGGCTCTGTTTCATGTGGGTGGATG 3652  
DB 766 GTGGAACCTCAACTCTTCTGTTCTCTTAGGGAAGTTCTAGACCCCGGTGGGTGGATA 825  
QY 3653 ACATTCAGTGTCTAAACGGATATCTCCATATAGGAGTGGCTCTCTGCCCCCATGGGAGC 3712  
DB 826 GAATCCAGTGTGGAAACTGATAGCTCTCTCTGGAGTGTCTCTGACCCCTTGGAAAT 885  
QY 3713 GAAGAATCTCCAGCCAGCAGACAGACCTGGATCACATGTGAAGATAGATAA----- 3766  
DB 886 ATAGTTCAATGTTCTTCAAGGAGGAGCCATATGTCTCTGTGCAAGGAGGAGACCCCAAGA 945  
QY 3767 -----GAGTGCCTGGAGGAGACA 3784  
DB 946 GCTGTCCAACCTACTGCCCTGCACAGACAAAGAGAGCTGCGACTCAGGGGAGGAGACA 1005  
QY 3785 CCGAGTGTCTGGAGAGTGGAGATCTGGCACGAGGCTCTGGGGCAGCAGTGTGTATG 3844  
DB 1006 GCCTGTGCTCAGGCGGAGTGGAGGTGGCACGCGGCTCTGGGGCAGCTGTGTGTATG 1065  
QY 3845 ACTCCTGGGACCTGGCGAGCGGGAAGTGTGTCTACAGCAGCTGGGCTGTGGCTCTGCTC 3904  
DB 1066 ACTCCTGGAGCCTGGCAGAGCCGAGGTGTGTCTACAGCAGCTGGGCTGTGGCCATGCC 1125  
QY 3905 TGGCTGCCCTCAGGAGCGCTTCTGTTGGCCAGGGAATGGAACCATCTGTTTGGATGACA 3964  
DB 1126 TGAAGCCCTTCGGGATCGGCATTTGGCCCTGGAAGTGGGAGCATCTGGCTGGACGAGG 1185  
QY 3965 TGGCGTGAAGAAATGAGTCAATTTCTATGGGACTGTACGCCAAACCTTGGGGACAGA 4024  
DB 1186 TGCAGTGCAGGGGAGGAGTCTCCCTGTGGGACTGTGTCTGGGGCCCTTGGGGGACAGA 1245  
QY 4025 GTGACTGTGGACACAGGAAGATGCTGGCGTGAAGTGTCTCTGG 4067  
DB 1246 GTGACTGCAAGCATGAGGAGGAGCGCTGGAGTGAATGCTCTGG 1288



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2003, 02:53:40 ; Search time 3980 seconds  
(without alignments)  
18832.340 Million cell updates/sec

Title: US-09-759-130B-379

Perfect score: 4628

Sequence: 1 gcgcgcctcgcgatctaga.....aaacggacgctggctgcac 4628

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	958.2	20.7	999	9	AL529439
2	578.8	12.5	582	12	BG570972
C 3	568.2	12.3	581	14	BQ776070
C 4	512.6	11.1	554	9	AL529438
5	491.4	10.6	940	9	AL550994
C 6	489.4	10.6	491	9	AI095706

C 7	453.8	9.8	457	9	AI738432
C 8	447.4	9.7	449	9	AI609329
C 9	446.2	9.6	451	10	AW082879
C 10	416.8	9.0	799	13	BI185454
C 11	414.4	9.0	429	9	AI278132
C 12	408.4	8.8	718	13	BI183796
C 13	407.8	8.8	411	9	AI090561
C 14	401.4	8.7	483	12	BF132072
C 15	392.8	8.5	606	14	BQ024006
C 16	390.8	8.2	392	14	W42792
C 17	359.4	8.0	371	14	N69291
C 18	359.4	8.0	385	14	W42886
C 19	366	7.9	822	12	BF792807
C 20	353.8	7.9	1100	9	AL540905
C 21	353.4	7.9	1108	9	AL544534
C 22	354	7.6	930	13	BI838644
C 23	324.2	7.0	705	17	BI8308
C 24	320	6.9	320	9	AA582301
C 25	310	6.7	723	13	BI836459
C 26	308.8	6.7	1153	9	AL540904
C 27	306.8	6.6	310	9	AL706449
C 28	304.8	6.6	484	14	N94111
C 29	301.8	6.5	644	13	BI009838
C 30	299.2	6.5	653	10	BE502724
C 31	296.8	6.4	616	13	BI009795
C 32	294.8	6.4	310	9	AI472927
C 33	291.6	6.3	585	14	BQ327934
C 34	290.4	6.3	408	9	AI288981
C 35	289.6	6.3	711	14	BQ378694
C 36	277.4	6.0	583	14	BQ368412
C 37	276.8	6.0	689	12	BF096204
C 38	275.2	5.9	569	10	AW631641
C 39	275.2	5.9	596	13	BI009802
C 40	269.8	5.8	574	14	BQ327932
C 41	268.2	5.8	452	17	AQ672871
C 42	267	5.8	833	9	AL574811
C 43	266.4	5.8	574	13	BI009286
C 44	265.4	5.7	560	13	BI009799
C 45	264.4	5.7	583	14	BQ327909

#### ALIGNMENTS

RESULT 1  
AL529439  
LOCUS  
DEFINITION  
AL529439 LTI\_NFL001\_NBC4 Homo sapiens CDNA clone CS0DD006Y005 5  
prime, mRNA sequence.  
ACCESSION  
AL529439  
VERSION  
AL529439.1 GI:12792932  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 999)  
AUTHORS  
Li W.B., Gruber C., Jessee J. and Polayes D.  
TITLE  
Full-length CDNA libraries and normalization  
JOURNAL  
Unpublished (2001)  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. 999  
/organism="Homo sapiens"  
/db\_xref="taxon:9605"  
/clone="CS0DD006Y005"  
/clone\_lib="LTI\_NFL001\_NBC4"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco kV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850; USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 220 a 236 c 318 g 224 t 1 others  
ORIGIN

Query Match 20.7%; Score 958.2; DB 9; Length 999;  
Best Local Similarity 99.3%; Pred. No. 6.8e-154;  
Matches 982; Conservative 1; Mismatches 4; Indels 2; Gaps 2;

QY 3205 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCAGGTGGTGTGTCAAAAGCTGGGCTGT 3264  
Db 1 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCAGGTGGTGTGTCAAAAGCTGGGCTGT 60

QY 3265 GGAGTGGCTTCAATCCACGGTCTCTGCTCACTTTGGGAGGGGTACAGGCCCATCTGG 3324  
Db 61 GGAGTGGCTTCAATGCCAGGTCTCTGCTCACTTTGGGAGGGGTACAGGCCCATCTGG 120

QY 3325 CTGGATGACCTGAACCTGCACAGGAACGGAGTCCCACTTTGGCAGTGCCTTTCCCGCGC 3384  
Db 121 CTGGATGACCTGAACCTGCACAGGAATGGAGTCCCACTTTGGCAGTGCCTTTCCCGCGC 180

QY 3385 TGGGGGAGCAGCTGACGAGCAAGAGGAGGAGCGAGGGTCTATCTGTCAAAATTCACA 3444  
Db 181 TGGGGGAGCAGCTGACGAGCAAGAGGAGGAGCGAGGGTCTATCTGTCAAAATTCACA 240

QY 3445 GCCTTGAGGCTTACAGTGAACCTGAACAGAGAGTGTCTGGGAGATTGGAAGTCTTC 3504  
Db 241 GCCTTGAGGCTTACAGTGAACCTGAACAGAGAGTGTCTGGGAGATTGGAAGTCTTC 300

QY 3505 TATAACGGGACCTGGGCGACGCTGGCAGGAGGAACATCACACAGCCATAGCAGGCATT 3564  
Db 301 TATAACGGGACCTGGGCGACGCTGGCAGGAGGAACATCACACAGCCATAGCAGGCATT 360

QY 3565 GTGTGAGGAGCTGGGCTGTGGGAGAGTGTGTGAGCTTCGCGCTTATCTAAG 3624  
Db 361 GTGTGAGGAGCTGGGCTGTGGGAGAGTGTGTGAGCTTCGCGCTTATCTAAG 420

QY 3625 ACAGGCTCTGTTTCATGTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATA 3684  
Db 421 ACAGGCTCTGTTTCATGTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATA 480

QY 3685 TGGCAGTGCCTGTCTGCCCCATGGGAGCAAGAATCTCCAGCCCCAGCAAGAGACCTGG 3744  
Db 481 TGGCAGTGCCTGTCTGCCCCATGGGAGCAAGAATCTCCAGCCCCAGCAAGAGACCTGG 540

QY 3745 ATACATCTGAGATAGAAATAGAGTCCGTGGAGGAGACACCGAGTCTCTGGGAGAGTG 3804  
Db 541 ATACATCTGAGATAGAAATAGAGTCCGTGGAGGAGACACCGAGTCTCTGGGAGAGTG 600

QY 3805 GAGATCTGGACGACGAGCTCTGGGGCACAGTGTGTGATGACCTCTGGGACCTGGCCGAG 3864  
Db 601 GAGATCTGGACGACGAGCTCTGGGGCACAGTGTGTGATGACCTCTGGGACCTGGCCGAG 660

QY 3865 GCGAAGTGTGTGACAGAGCTGGGCTGTGGCTCTGCTCTGCTGCTGAGGAGGAGCT 3924  
Db 661 GCGAAGTGTGTGACAGAGCTGGGCTGTGGCTCTGCTCTGCTGCTGAGGAGGAGCT 720

QY 3925 TCGTTTGGCCAGGAATGGAACCATCTGTTGATGACATGGGTGCAAGGAAATGAG 3984  
Db 721 TCGTTTGGCCAGGAATGGAACCATCTGTTGATGACATGGGTGCAAGGAAATGAG 780

QY 3985 TCATTTCTATGGGACTGTACAGCCAAACCTCTGGGACAGAGTGAAGTGTGACACAGGAA 4044  
Db 781 TCATTTCTATGGGACTGTACAGCCAAACCTCTGGGACAGAGTGAAGTGTGACACAGGAA 840

QY 4045 GATCTGGCGTGGAGGTGCTCTGGACAGTCGCTGAATCACTGAATGCTCTCAGGTCAAT 4104  
Db 841 GATCTGGCGTGGAGGTGCTCTGGACAGTCGCTGAATCACTGAATGCTCTCAGGTCAAT 900

QY 4105 TTAGCACTTATTTATCCAGTATCTTTGGGCTCTCTCCCTGGTCTCTGTTATTTCTATT 4164  
Db 901 TTAGCACTTATTTATCCAGTATCTTTGGGCTCTCTCCCTGGTCTCTGTTATTTCTATT 958

QY 4165 CTCACGTGGTCCCGAGTTCAGAAACAAA 4193  
Db 959 CTCACGTGGTCCCGAGTTCAGAAACAAA 987

RESULT 2  
BG570972 582 bp mRNA linear EST 10-APR-2001  
LOCUS 602591546F1 NIH\_MGC\_79 Homo sapiens cDNA clone IMAGE:4714115 5',  
DEFINITION mRNA sequence.  
ACCESSION BG570972  
VERSION BG570972.1 GI:13578625  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 582)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: CLONTECH Laboratories, Inc.  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LICM1558 row: k column: 12  
High quality sequence stop: 576.  
Location/Qualifiers

FEATURES  
source

1..582  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4714115"  
/clone\_lib="NIH\_MGC\_79"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);  
Site\_1: SfiI (ggcgctcgcc); Site\_2: SfiI (ggccattatggcc  
); 5' and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-ATTCTAGAGCCGAGGCGGCGACATG-dT(30)BN-3'  
sequence: 5'-ATTCAGAGCCGAGGCGGCGGCGACATG-dT(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
Library.

BASE COUNT 126 a 140 c 198 g 118 t  
ORIGIN

Query Match 12.5%; Score 578.8; DB 12; Length 582;  
Best Local Similarity 99.7%; Pred. No. 4.6e-89;  
Matches 580; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3351 GGAGTCCCCTTGTGGCAGTGCCTTCCCGCGGCTGGGGCAGCAGACTGCAGGCACAA 3410  
Db 1 GGAGTCCCCTTGTGGCAGTGCCTTCCCGCGGCTGGGGCAGCAGACTGCAGGCACAA 60

QY 3411 GGAGGACGAGGGGTCTATCTGTCTCAGAAATTCAGCCTTTCAGCTCTACAGTGAAGTGA 3470  
Db 61 GGAGGACGAGGGGTCTATCTGTCTCAGAAATTCAGCCTTTCAGCTCTACAGTGAAGTGA 120

Qy	3471	AACAGACAGCTGCTCTGGGAGATTGGAAGCTTCTTATAACGGGACCTGGGCGACGCTCGG	3530
Db	121	AACAGACAGCTGCTCTGGGAGATTGGAAGCTTCTTATAACGGGACCTGGGCGACGCTCGG	180
Qy	3531	CAGGAGGAACATACCAACAGCATAGCAGCATTGTGTGCAGGACGCTGGCTGTGGGA	3590
Db	181	CAGGAGGAACATACCAACAGCATAGCAGCATTGTGTGCAGGACGCTGGCTGTGGGA	240
Qy	3591	GAATGGAGTTGTACGCTCGCCCTTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGA	3650
Db	241	GAATGGAGTTGTACGCTCGCCCTTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGA	300
Qy	3651	TGACATTCACTGTCCTAAACGGCATATCTCCATATGCGAGTGCCTGTCTGCCCATGGA	3710
Db	301	TGACATTCACTGTCCTAAACGGCATATCTCCATATGCGAGTGCCTGTCTGCCCATGGA	360
Qy	3711	CGAAGAATCTCCAGCCCCAGCAGAGACCTGGATCACATGTGAAGATAGAATAAGAGT	3770
Db	361	CGAAGAATCTCCAGCCCCAGCAGAGACCTGGATCACATGTGAAGATAGAATAAGAGT	420
Qy	3771	CGCTGGAGGAGACACCGAGTGCTCTGGAGAGTGGAGATCTGGCACGAGGCTCCTGGGG	3830
Db	421	CGCTGGAGGAGACACCGAGTGCTCTGGAGAGTGGAGATCTGGCACGAGGCTCCTGGGG	480
Qy	3831	CACAGTCTGTGATGACATCCCTGGGACCTGGGCCGAGCGGAAGTGTGTGTCAGCAGCTGGG	3890
Db	481	CACAGTCTGTGATGACATCCCTGGGACCTGGGCCGAGCGGAAGTGTGTGTCAGCAGCTGGG	540
Qy	3891	CTGTGGCTCTGCTCTGGCTGCGCTTGAGGAGCGCTTCGTTTGG	3932
Db	541	CTGTGGCTCTGCTCTGGCTGCGCTTGAGGAGCGCTTCGTTATGG	582

RESULT 3	BQ776070/c	BQ776070	581 bp	mRNA	linear	EST 26-JUL-2002
LOCUS						
DEFINITION		UI-H-FHO-bcl-f-19-0-UI.s1	NCI_CGAP_FHO	Homo sapiens	cdna clone	
DEFINITION		UI-H-FHO-bcl-f-19-0-UI.3'	mRNA sequence.			

FEATURES	Location/Qualifiers
POLYMERIES.	
source	1. .581
	/organism="Homo sapiens"
	/db_xref="taxon.9606"
	/clone="UI-H-FHO-bcl-f-19-0-UI"
	/clone_lib="NCI_CGAP_FHO"
	/tissue_type="Human Chondrosarcoma Cell Line"
	/dev_stage="Adult"
	/lab_host="Dr10B (Life Technologies)"
	/notes="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site.1: EcoR I; Site.2: Not I;

RESULT 4	AL529438/c	AL529438	554 bp	mrna	linear	EST 13-FEB-2001
LOCUS	AL529438	AL529438	LTI_NFL001_NBC4	Homo sapiens	cdna clone	CSODD006Y005 3
DEFINITION	prime, mRNA sequence.					
ACCESSION	AL529438	AL529438				
VERSION	AL529438.1	GI:12792931				
KEYWORDS	EST.					

```

SOURCE          human.
ORGANISM        Homo sapiens
REFERENCE       Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 554)
JOURNAL        Full-length cDNA libraries and normalization
COMMENT        Unpublished (2001)
CONTACT        Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES       1..554
SOURCE         Location/Qualifiers
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="CSDD006Y005"
               /clone_lib="LTI_NFL001_NBC4"
               /sex="male"
               /tissue_type="neuroblastoma cells"
               /lab_host="DH10B"
               /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
               was primed with a NotI-oligo(dT) primer. Five prime end
               enriched, double-stranded cDNA was digested with Not I and
               cloned into the Not I and Eco RV sites of the pCMVSPORT 6
               vector. Library was normalized. Library was constructed
               by Life Technologies. Contact : Feng Liang Life
               Technologies, a division of Invitrogen 9800 Medical Center
               Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
               8371 Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com"
BASE COUNT     132 a 166 c 127 g 110 t 19 others
ORIGIN
Query Match    11.1%; Score 512.6; DB 9; Length 554;
Best Local Similarity 94.8%; Pred. No. 9.1e-78;
Matches 525; Conservative 9; Mismatches 19; Indels 1; Gaps 1;

QY 3592 AATGGAGTCTCAGCCTCGCCCTTTA-TCTAAGACAGGCTCTGGTTTCATGTGGGTGA 3650
DB 554 AATGGAGTATACGCTCGCCCTTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGA 495
QY 3651 TGACATTGAGTCTTAAAGCGATATCTCCATATGGCAGTGCCTCTCGCCCATGGGA 3710
DB 494 TRACATTGAGTCTTAAAGCGATATCTCCATATGGCAGTGCCTCTCGCCCATGGGA 435
QY 3711 GCGAAGATCTCAGCCACAGACAGACCTGGATACATGTGAAGATAGATAAGACT 3770
DB 434 GCGAAGATCTCAGCCACAGACAGACCTGGATACATGTGAAGATAGATAAGACT 375
QY 3771 GCGTGGAGGACACACCGAGTCTCTGGAGAGTGGAGATCTGGCAGCGAGTCTCTGGGG 3830
DB 374 GCGTGGAGGACACACCGAGTCTCTGGAGAGTGGAGATCTGGCAGCGAGTCTCTGGGG 315
QY 3831 CACAGTGTGTGATGACTCTGGACCTTGGCCGAGCGGAGTGTGTGTCAGCAGCTGGG 3890
DB 314 CACAGTGTGTGATGACTCTGGACCTTGGCCGAGCGGAGTGTGTGTCAGCAGCTGGG 255
QY 3891 CTGTGGCTCTGCTCTGCTGCTGCTGAGGACGCTTCTGTTGGCCAGGAACTGGAACCAT 3950
DB 254 CTGTGGCTCTGCTCTGCTGCTGAGGACGCTTCTGTTGGCCAGGAACTGGAACCAT 195
QY 3951 CTGTTGGATGACATGCGGTGCAAGAAATGAGTCAATTTCTATGGAGTGTCAACGCCAA 4010
DB 194 CTGTTGGATGACATGCGGTGCAAGAAATGAGTCAATTTCTATGGAGTGTCAACGCCAA 135
QY 4011 ACCCTGGGACAGAGTCACTGTGGACACAAGGAGATGCTGGCGTGAAGTCTCTGGACA 4070
DB 134 ACCCTGGGACAGAGTCACTGTGGACACAAGGAGATGCTGGCGTGAAGTCTCTGGACA 75
QY 4071 GTCGCTGAAATCACTGAATGCTCTCAGGTCAATTTAGCAGTATATTTTCCAGTATCTT 4130
DB 74 GTCGCTGAAATCACTGAATGCTCTCAGGTCAATTTAGCAGTATATTTTCCAGTATCTT 15

```

```

QY 4131 TGGGCTCCTCTCTCC 4144
DB 14 TGGGCTCCTCTCTCC 1

RESULT 5
AL550994 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI066YE07 5
LOCUS prime, mRNA sequence.
DEFINITION AL550994 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI066YE07 5
ACCESSION AL550994
VERSION AL550994.1 GI:12888516
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 940)
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
CONTACT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES       1..940
SOURCE         Location/Qualifiers
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="CSODI066YE07"
               /clone_lib="LTI_NFL006_PL2"
               /tissue_type="placenta"
               /note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
               was primed with a NotI-oligo(dT) primer. Five prime end
               enriched, double-stranded cDNA was digested with Not I and
               cloned into the Not I and Eco RV sites of the pCMVSPORT 6
               vector. Library was normalized. Library was constructed by
               Life Technologies. Contact : Feng Liang Life Technologies,
               a division of Invitrogen 9800 Medical Center Drive
               Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
               Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com"
BASE COUNT     228 a 212 c 279 g 218 t 3 others
ORIGIN
Query Match    10.6%; Score 491.4; DB 9; Length 940;
Best Local Similarity 71.3%; Pred. No. 3e-74;
Matches 645; Conservative 2; Mismatches 258; Indels 0; Gaps 0;

QY 2599 TGGCGCGAAGTCTCAGTGTGAAGGAGTGAACCTCACCTTCATTTGCCCATTTGT 2658
DB 9 TGGCGTGAAGATTCAGTGTGAGGAGCATGAGTCCCATCTTCTACTCTGCYCACTAGCA 68
QY 2659 CAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTTCCCGAT 2718
DB 69 CCCCCCCACAGAGGAACCTGTAGCCACAGCAGGAGTGTGGAGTAGTGTCTCAAGATAC 128
QY 2719 ACAGATGTCGCACTTGTGAATGGCAATCCCAATGTCAGCGGCAAGTGGAGATCAACGTG 2778
DB 129 ACAGAAATTCGTTGTTGAATGGCAAGCCCGTGTGAGGGCAGAGTGGAGTCTAAACG 188
QY 2779 CTTGGACACTGGGGTCACTGTGTACACCCACTGGAGCCAGAGAGTCCCGTGTCTTA 2838
DB 189 CTTGGTGCCTGGGGTCCCTCTGTACTCTCACTGGACATAGAGATGCCCATGTTCT 248
QY 2839 TGCAGACAGCTCAGCTGTGGGACTGCTCTCAACACAGGAGGAAATATATTGGAGAA 2898
DB 249 TGCAGACAGCTTAAATGTGGAGTTGCCCTTCTACCCAGGAGGACGCTTTTGGAAA 308
QY 2899 AGNAGTGTTCGTGTGGGACAGAGTTTCAATGTTAGGGAATGAGTCACTTCTGGAT 2958
DB 309 GGAATGGTCAAGATCTGGAGGCATATGTTTCACTGCGGACTGGAGGCACATGGGA 368

```



Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.  
Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 600 Std Error: 0.00  
Seq primer: -40UP from Gibco.

FEATURES

Location/Qualifiers  
1. 457  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2391130"  
/clone\_lib="NCI-CGAP\_Col6"  
/tissue\_type="colon tumor, RER"  
/lab\_host="DH10B"  
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Plasmid DNA from the normalized library NCI-CGAP\_Col6 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneIDs 1057416-1061255, and 1144584-1145351).  
Subtraction by Bento Soares and M. Fatima Bonaudo. "  
BASE COUNT 119 a 94 c 100 g 144 t

Query Match 9.8%; Score 453.8; DB 9; Length 457;  
Best Local Similarity 99.6%; Pred. No. 1.1e-67;  
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4128 CTTTGGGCTCTCTCCTGGTCTGTTTATTTCTATTTCTACAGTGGTCCGAGTTCAGAA 4187  
|||||  
DB 457 CTTTGGGCTCTCTCCTGGTCTGTTTATTTCTATTTCTACAGTGGTCCGAGTTCAGAA 398  
|||||  
QY 4188 ACAAACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGAGGAGAATTT 4247  
|||||  
DB 397 ACAAACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGAGGAGAATTT 338  
|||||  
QY 4248 ATTCCATGATGGAGACCTGCTCAAGAGAGAGGACCCACATGGGACAAAGACCTCAGA 4307  
|||||  
DB 337 ATTCCATGATGGAGACCTGCTCAAGAGAGAGGACCCACATGGGACAAAGACCTCAGA 278  
|||||  
QY 4308 TGACACCCCAACCATGGTTGTAAGATGCTAGCGACATCGTCTGGGAGTCTTCC 4367  
|||||  
DB 277 TGACACCCCAACCATGGTTGTAAGATGCTAGCGACATCGTCTGGGAGTCTTCC 218  
|||||  
QY 4368 TGCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAAT 4427  
|||||  
DB 217 TGCTCTGAAGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAAT 158  
|||||  
QY 4428 ATCTTTGAAGGAGACAAACACTTTTAAATGAATGAAGAGAGTCAAGTTGCCCTATGA 4487  
|||||  
DB 157 ATCTTTGAAGGAGACAAACACTTTTAAATGAATGAAGAGAGTCAAGTTGCCCTATGA 98  
|||||  
QY 4488 AAACCTGTCCAAATAACATTTCTTTGAACAATAGGACAAAGCTAAATTTGATAAAGACTGG 4547  
|||||  
DB 97 AAACCTGTCCAAATAACATTTCTTTGAACAATAGGACAAAGCTAAATTTGATAAAGACTGG 38  
|||||  
QY 4548 TGATAATAAAATGAATTTATGATATCATCTGTTAAA 4584  
|||||  
DB 37 TGATAATAAAATGAATTTATGATATCATCTGTTAAA 1

RESULT 8.  
AI609329/c  
LOCUS  
DEFINITION  
449 bp mRNA linear EST 14-MAY-1999  
tw30b02.x1 NCI-CGAP\_Ov35 Homo sapiens cDNA clone IMAGE:2261163 3',  
mRNA sequence.

ACCESSION AI609329  
VERSION AI609329.1 GI:4618496  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 449)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP);  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 631 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 415  
POLYA-No.

FEATURES

Location/Qualifiers  
1. 449  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2261163"  
/clone\_lib="NCI-CGAP\_Ov35"  
/tissue\_type="tumor, 5 pooled (see description)"  
/lab\_host="DH10B"  
/note="Organ: ovary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; This library represents the normalized  
version of NCI-CGAP\_Ov23. Cloned unidirectionally.  
Primer: Oligo dt. Average insert size 0.86 kb. Tumor  
types include: mixed Mullerian tumor, papillary serous,  
clear cell, spindle cell. All are primary tumors,  
metastasis positive. Constructed by Life Technologies."  
BASE COUNT 116 a 94 c 97 g 142 t

Query Match 9.7%; Score 447.4; DB 9; Length 449;  
Best Local Similarity 99.8%; Pred. No. 1.3e-66;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4133 GGCTCTCTCTCTGTTCTGTTTATTTCTATTTCTCACGTGGTCCGAGTTCAGAAACAAA 4192  
|||||  
DB 449 GGCTCTCTCTCTGTTCTGTTTATTTCTATTTCTCACGTGGTCCGAGTTCAGAAACAAA 390  
|||||  
QY 4193 AACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCAGAGAGATTTATTC 4252  
|||||  
DB 389 AACATCTGCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCAGAGAGATTTATTC 330  
|||||  
QY 4253 ATGAGATGGAGACCTGCCTCAGAGAGAGGACCCACATGGGACAAAGACCTCAGATGACA 4312  
|||||  
DB 329 ATGAGATGGAGACCTGCCTCAGAGAGAGGACCCACATGGGACAAAGACCTCAGATGACA 270  
|||||  
QY 4313 CCCCACCAACCATGTTGTGAAGATGCTAGGCACACATCGCTGTTGGGAGTTCCTTCCTGCCT 4372  
|||||  
DB 269 CCCCACCAACCATGTTGTGAAGATGCTAGGCACACATCGCTGTTGGGAGTTCCTTCCTGCCT 210  
|||||  
QY 4373 CTGAAGCCCAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 4432  
|||||  
DB 209 CTGAAGCCCAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTT 150  
|||||  
QY 4433 TGAAGGAGACACAACTTTTAAATGAATAAGAGAGAGTTCAGTTCGCCCTATGGAAACT 4492  
|||||  
DB 149 TGAAGGAGACACAACTTTTAAATGAATAAGAGAGAGTTCAGTTCGCCCTATGGAAACT 90  
|||||  
QY 4493 TGTCCAAATAAATCTTTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACTGGTGATA 4552

```
|||||
Db 89 TGTCACAAATACATTTCTTGAACAATAGGAGAACAGCTAAATTGATAAAGACTGGTGATA 30
|||||
QY 4553 ATAAAAATTCGAATATATATATCACTGTT 4581
|||||
Db 29 ATAAAAATTCGAATATATATATCACTGTT 1
|||||

RESULT 9
AW082879/c AW082879 451 bp mRNA linear EST 14-OCT-1999
LOCUS
DEFINITION
IMAGE: x571g11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
ACCESSION
VERSION
AW082879
SOURCE
EST.
GI:6038031
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 451)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
FEATURES
source
1..451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2581796"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 115 a 91 c 96 g 149 t
ORIGIN
Query Match 9.6%; Score 446.2; DB 10; Length 451;
Best Local Similarity 99.3%; Pred. No. 2.1e-66;
Matches 448; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4138 CTTCTCTGGTCTGTTATTCTATTTCTACGTTGGTCCGAGTTCAGAACAAACAT 4197
|||||
Db 451 CTTCTCTGGTCTGTTATTCTATTTCTACGTTGGTCCGAGTTCAGAACAAACAT 392
|||||
QY 4198 CTGCCCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGAGGAGAATTTATTCATGAG 4257
|||||
Db 391 CTGCCCCCTCAGAGTTTCAACAGAGAGGGGTTCTCTCGAGGAGAATTTATTCATGAG 332
|||||
QY 4258 ATGAGACCTCCCTCAGAGAGAGGCCACATGGGACAGAACCTCAGATGACACCCCC 4317
|||||
Db 331 ATGAGACCTCCCTCAGAGAGAGGCCACATGGGACAGAACCTCAGATGACACCCCC 272
|||||
QY 4318 AACCATGGTCTGAAGTCTAGGCACACATCGCTGTGGGAGTCTCTCGCCCTCTGAA 4377
|||||
Db 271 AACCATGGTCTGAAGTCTAGGCACACATCGCTGTGGGAGTCTCTCGCCCTCTGAA 212
|||||
QY 4378 GCCACAAATGACTTTAGACTTCCAGGCTCACCAGATCAACCTCTAAATATCTTTGAAG 4437
|||||
```

```
|||||
Db 211 GCCACAAATGACTTTAGACTTCCAGGCTCCACCAGATCAACCTCTAAATATCTTTGAAG 152
|||||
QY 4438 GAGACAACAACCTTTTAATGAATAAGAGAGAGTCAAGTTGCCCTATGGAAACCTTGCC 4497
|||||
Db 151 GAGACAACAACCTTTTAATGAATAAGAGAGAGTCAAGTTGCCCTATGGAAACCTTGCC 92
|||||
QY 4498 AAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACTGGTGATAATAA 4557
|||||
Db 91 AAATAACATTTCTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACTGGTGATAATAA 32
|||||
QY 4558 AATTGAATTTATGATATCACTGTTTAAAAAAA 4588
|||||
Db 31 AATTGAATTTATGATATCACTGTTTAAAAAAA 1
|||||

RESULT 10
BI185454 799 bp mRNA linear EST 10-JUL-2001
LOCUS
DEFINITION
UNL-P-FN-cw-d-01-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-cw-d-01-0-UNL 3', mRNA sequence.
ACCESSION
VERSION
BI185454.1 GI:14659863
KEYWORDS
EST.
SOURCE
pig.
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 799)
Caetano, A.R., Johnson, R.K. and Pomp, D.
Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
JOURNAL
Unpublished (2001)
COMMENT
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dt track not found, Not I site shown in beginning of sequence
is likely internal to the message.
Seq primer: M13 -29
POLYA-No.
FEATURES
source
Location/Qualifiers
1..799
/organism="Sus scrofa"
/strain="University of Nebraska, Lincoln Swine Selection
Lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-cw-d-01-0-UNL"
/clone_lib="UNL-P-FN"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UNL-P-FN
library is a normalized library representing porcine
ovarian follicles, ranging between 2.0 to 10.0 mm in
diameter, collected during 7 days of the follicular phase
of the pig estrous cycle. This library was derived from
the library UNL-P-F2. The tag is a string of 5-6
nucleotides present between the Not I site and the
oligo-dt track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996."
BASE COUNT 206 a 174 c 177 t 1 others
ORIGIN
Query Match 9.0%; Score 416.8; DB 13; Length 799;
Best Local Similarity 71.1%; Pred. No. 1.7e-61;
Matches 564; Conservative 0; Mismatches 228; Indels 1; Gaps 1;
```





REFERENCE	1 (bases 1 to 606)
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)
	Contact: Robert Strausberg, Ph.D.
	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>
	Tissue Procurement: Dr. Steven Brown
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
	cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

## FEATURES

Location/Qualifiers

```

1. 606
location/Qualifiers
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="UI-1-BB1p-auu-d-07-0-UI"
  /clone_lib="NCI_CGAP_P16"
  /tissue_type="Placenta"
  /dev_stage="Full Term"
  /lab_host="DH10B (Life Technologies)"
  /note="Organ: Placenta; Vector: pT73-Pac (Pharmacia) with
a modified polylinker; Site_1: Ecor I; Site_2: Not I;
NCI_CGAP_P16 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are GA, AGGAA. For
additional information, contact: Bento Soares,
bento-soares@uiowa.edu
TAG_LIB=UI-1-BB1p
TAG_TISSUE=placenta human full term
TAG_SEQ=AGGAA"
127 g 229 t

```

Query Match	8.5%;	Score 392.8;	DB 14;	Length 606;
Best Local Similarity	98.3%;	Pred. No. 2.4e-57;		
Matches 397;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
QY	4196	ATCTGCCCCCTCAGAGTTTCAACAGAGGAGGGGTCTCTCGAGGGAGAAATTTATTCATG	42555	
Db				
Db	404	AGCATCCATTCTCAGTTTCAACAGAGGAGGGGTCTCTCGAGGAGAAATTTATTCATG	345	
QY	4256	AGATGGAGAGCTGCCTCAGAGAGAGAGACCCACATGGGACAGAAACCTCAGATGACACCC	4315	
Db				
Db	344	AGATGGAGAGCTGCCTCAAGAGAGAGAGACCCACATGGGACAGAAACCTCAGATGACACCC	285	
QY	4316	CCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTGTGGGAGTTCTTCCTGCCCTCTG	4375	
Db				
Db	284	CCAACCATGGTTGTGAAGATGCTAGCGACACATCGCTGTGTGGGAGTTCTTCCTGCCCTCTG	225	
QY	4376	AGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGA	4435	
Db				
Db	224	AGCCACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGA	165	
QY	4436	AGGAGACAAACACTTTTAAATGAATAAGAGGAGTCAAGTTGCCCTATGGAAAACCTGT	4495	
Db				
Db	164	AGGAGACAAACACTTTTAAATGAATAAGAGGAGTCAAGTTGCCCTATGGAAAACCTGT	105	
QY	4496	CCAAAATCAACTTCTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACGTGGTGATAATA	4555	
Db				
Db	104	CCAATAACACTTCTTGAACAATAGGAGAACAGCTAAATTTGATAAAGACGTGGTGATAATA	45	
QY	4556	AAATTTGAATATGTATATACCTGTTAAAAAIAAAAAAAAAAAAAA	4599	
Db				
Db	44	AAATTTGAATATGTATATACCTGTTAAAAAIAAAAAAAAAAAAAA	1	

Search completed: May 12, 2003, 06:24:01  
Job time : 4026 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2003, 02:56:30 ; Search time 614 Seconds  
(without alignments)  
9366.838 Million cell updates/sec

Title: US-09-759-130B-379

Perfect score: 4628

Sequence: 1 ggcggcgtcgatctaga.....aaacgagcgtgggtcgac 4628

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4628	100.0	4628	9	US-10-042-431-9
2	4628	100.0	4628	9	US-09-759-130B-379
3	4359	94.2	4359	9	US-10-042-431-10
4	4359	94.2	4359	9	US-09-759-130B-380
5	886.8	19.2	4308	9	US-10-042-431-79
6	886.8	19.2	4308	9	US-09-759-130B-449
7	447.4	9.7	449	10	US-09-867-701-10340
8	434.6	9.4	473	9	US-09-918-995-32438
9	213	4.6	599	9	US-10-040-739-138
10	199.2	4.3	461	9	US-09-796-692-269
11	199.2	4.3	461	9	US-09-796-692-7354
12	199.2	4.3	461	9	US-10-040-862-269
13	199.2	4.3	461	9	US-10-040-862-7354
14	197.6	4.3	2562	9	US-09-147-947-5
15	193.2	4.2	2614	9	US-09-147-947-3
16	173.4	3.7	444	10	US-09-925-301-696
17	169.8	3.7	438	9	US-09-918-995-35432
18	149.6	3.2	4344	10	US-09-917-800A-1712
19	143.2	3.1	697	10	US-09-925-297-237

20	134.4	2.9	713	9	US-10-040-739-1356	Sequence 1356, App
21	123.4	2.7	1686	9	US-09-905-291A-147	Sequence 147, App
22	123.4	2.7	1686	9	US-09-902-853-147	Sequence 147, App
23	123.4	2.7	1686	9	US-09-907-824-147	Sequence 147, App
24	123.4	2.7	1686	9	US-09-907-841-147	Sequence 147, App
25	123.4	2.7	1686	9	US-09-904-011-147	Sequence 147, App
26	123.4	2.7	1686	9	US-09-906-742-147	Sequence 147, App
27	123.4	2.7	1686	9	US-09-906-838-147	Sequence 147, App
28	123.4	2.7	1686	9	US-09-907-613-147	Sequence 147, App
29	123.4	2.7	1686	9	US-09-907-942-147	Sequence 147, App
30	123.4	2.7	1686	9	US-10-227-884-113	Sequence 113, App
31	123.4	2.7	1686	9	US-09-904-820-147	Sequence 147, App
32	123.4	2.7	1686	9	US-09-904-859-147	Sequence 147, App
33	123.4	2.7	1686	9	US-09-909-204-147	Sequence 147, App
34	123.4	2.7	1686	9	US-10-230-163-113	Sequence 113, App
35	123.4	2.7	1686	9	US-09-904-786-147	Sequence 147, App
36	123.4	2.7	1686	9	US-09-906-646-147	Sequence 147, App
37	123.4	2.7	1686	9	US-09-906-700-147	Sequence 147, App
38	123.4	2.7	1686	9	US-09-902-903-147	Sequence 147, App
39	123.4	2.7	1686	9	US-09-903-749A-147	Sequence 147, App
40	123.4	2.7	1686	9	US-09-903-786-147	Sequence 147, App
41	123.4	2.7	1686	9	US-10-218-631-113	Sequence 113, App
42	123.4	2.7	1686	9	US-10-230-338-113	Sequence 113, App
43	123.4	2.7	1686	9	US-09-902-736-147	Sequence 147, App
44	123.4	2.7	1686	9	US-09-904-119-147	Sequence 147, App
45	123.4	2.7	1686	9	US-09-904-956-147	Sequence 147, App

ALIGNMENTS

RESULT 1

US-10-042-431-9  
; Sequence 9, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: FRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042.431  
; PRIOR FILING DATE: 2001-10-25  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 4628  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-042-431-9

Query Match	100.0%	Score 4628	DB 9	Length 4628
Best Local Similarity	100.0%	Pred. No. 0	Mismatches 0	Indels 0
Matches 4628	Conservative	0	0	Gaps 0
Qy	1	CGCGCGCTCGGATCTAGAACTAGTATGATGCTGCTCAAACTCTGGCATATTGAT	60	
Db	1	CGCGCGCTCGGATCTAGAACTAGTATGATGCTGCTCAAACTCTGGCATATTGAT	60	
Qy	61	TTTGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGGTAACTTGCATCTGCTC	120	
Db	61	TTTGAAGATGCTGCTGTCATCAGAACCTTTTCTCTGCTGGTAACTTGCATCTGCTC	120	
Qy	121	CTGAATTCCTGCTTCTCATCAGAGTTTAAATGGAACAGATTTGGAGTTGAGGCTGTC	180	
Db	121	CTGAATTCCTGCTTCTCATCAGAGTTTAAATGGAACAGATTTGGAGTTGAGGCTGTC	180	

QY 181 AATGGACGGTCCCTGCTCTGGACAGTGGAGTGAATTCAGGACAGTGGGGACT 240  
DB 181 AATGGACGGTCCCTGCTCTGGACAGTGGAGTGAATTCAGGACAGTGGGGACT 240  
QY 241 GTGTGTGATGATGGGTGGACACTACTGCTCAACTGCTGTCGTGCAACAGCTTGGATGT 300  
DB 241 GTGTGTGATGATGGGTGGACACTACTGCTCAACTGCTGTCGTGCAACAGCTTGGATGT 300  
QY 301 CCATTTTCTTTCCCATGTTTCGTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGG 360  
DB 301 CCATTTTCTTTCCCATGTTTCGTTTGGACAAGCCGTGACTAGACATGGAAAAATTTGG 360  
QY 361 CTGTGATGATGTTTCCCTGTTTATGGAATGAGTCACTCTCTGGGAATGTCACACCGGGAA 420  
DB 361 CTGTGATGATGTTTCCCTGTTTATGGAATGAGTCACTCTCTGGGAATGTCACACCGGGAA 420  
QY 421 TGGGGAGCCATAACTGTTTATCATGAGAAGATGTTGGTGTGAACCTGTTATGGTGAAGCC 480  
DB 421 TGGGGAGCCATAACTGTTTATCATGAGAAGATGTTGGTGTGAACCTGTTATGGTGAAGCC 480  
QY 481 AATCTGGGTTTGAAGCTAGTGGATGGAACAACTCCTGTTCAAGGAGAGTGGAGGTGAAA 540  
DB 481 AATCTGGGTTTGAAGCTAGTGGATGGAACAACTCCTGTTCAAGGAGAGTGGAGGTGAAA 540  
QY 541 TTCCAAAGAAAGTGGGGGACTATATGTGATGATGAGTGGTGGAACTTGAATACTGTCGCCGTG 600  
DB 541 TTCCAAAGAAAGTGGGGGACTATATGTGATGATGAGTGGTGGAACTTGAATACTGTCGCCGTG 600  
QY 601 GTGTGACGCAACTAGGATGCTCATCTCTTTTATTTCTCTGGAGTCTGTTAATAGCCCT 660  
DB 601 GTGTGACGCAACTAGGATGCTCATCTCTTTTATTTCTCTGGAGTCTGTTAATAGCCCT 660  
QY 661 GCTGTATTCGCCGCCATTTGGCTGSGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720  
DB 661 GCTGTATTCGCCGCCATTTGGCTGSGATGACATTTTATGCCAGGGGAATGAGTTGGCACTC 720  
QY 721 TGAATTTGACAGACATCGTGGATGGGAATCATGACTGACGTACACATGAGGATGTCACA 780  
DB 721 TGAATTTGACAGACATCGTGGATGGGAATCATGACTGACGTACACATGAGGATGTCACA 780  
QY 781 TTAACCTGTTATCATAGTAGTAGTCTTTGAACAAAGCTTTGAGTGGAACTAACCCCTGT 840  
DB 781 TTAACCTGTTATCATAGTAGTAGTCTTTGAACAAAGCTTTGAGTGGAACTAACCCCTGT 840  
QY 841 ATGGGAGAGTAGAGTGAATAATCCAAGGAAGTGGGGACCGTATGCCACCAATAAGTGG 900  
DB 841 ATGGGAGAGTAGAGTGAATAATCCAAGGAAGTGGGGACCGTATGCCACCAATAAGTGG 900  
QY 901 AACAAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACCTC 960  
DB 901 AACAAATGCTGCAGCTGATGTCGTATGCAAGCAGTTGGGATGTGGAACCGCACTTCACCTC 960  
QY 961 GCTGGCTTGCTCATTTGCACTCAGGCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020  
DB 961 GCTGGCTTGCTCATTTGCACTCAGGCTGATGTTGTATGGCTTGATGGTGTCTCCTGC 1020  
QY 1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCGGGAACCGCTCAATTTTGAAGTGT 1080  
DB 1021 TCCGGTAATGAATCTTTTCTTTGGGACTGCAGACATTCGGGAACCGCTCAATTTTGAAGTGT 1080  
QY 1081 CTTTCATCAAAACCATGCTGCTGATCTGCTCAGATGGAGCAGATTTGGNACTGGCAGCTA 1140  
DB 1081 CTTTCATCAAAACCATGCTGCTGATCTGCTCAGATGGAGCAGATTTGGNACTGGCAGCTA 1140  
QY 1141 GCAGATGGAAGTAAACATTTCTCAGGAGAGTAGAGTGGAGATTCATGAACAGTGGTGG 1200  
DB 1141 GCAGATGGAAGTAAACATTTCTCAGGAGAGTAGAGTGGAGATTCATGAACAGTGGTGG 1200  
QY 1201 ACAATATGTGACCAGAACTGGAGATGAACAAGCCCTTGTGTTGTGAAGCAGCTAGGA 1260  
DB 1201 ACAATATGTGACCAGAACTGGAGATGAACAAGCCCTTGTGTTGTGAAGCAGCTAGGA 1260  
QY 1261 TGTCCGTTGACGCTCTTTGGCAGTCGTCGTGCTTAACCTAGTAATGAGCTAGACATTT 1320

DB 1261 TGTCCGTTGACGCTCTTTGGCAGTCGTCGTGCTAAACCTAGTAATGAACTAGACATTT 1320  
QY 1321 TCGATAAACAGCATATCTTTGCACTGGGAATGAGTCACTCTCTGGGACTGCACATATGAT 1380  
DB 1321 TCGATAAACAGCATATCTTTGCACTGGGAATGAGTCACTCTCTGGGACTGCACATATGAT 1380  
QY 1381 GGAAAGCAAAACGAAACATGCTTCCGAAGATCAGATGCTGAGATTAATTTGTTCTGATAAG 1440  
DB 1381 GGAAAGCAAAACGAAACATGCTTCCGAAGATCAGATGCTGAGATTAATTTGTTCTGATAAG 1440  
QY 1441 GCAGATCTGGACCTTAAGGCTTGTGGGGCTCATAGCCCTGTTATGSGAGATGGAGGTG 1500  
DB 1441 GCAGATCTGGACCTTAAGGCTTGTGGGGCTCATAGCCCTGTTATGSGAGATGGAGGTG 1500  
QY 1501 AATATCCAAAGGAGTGGGGGACTGTGCTCATGACAGATGAGCACAAAGGAATGCAGCT 1560  
DB 1501 AATATCCAAAGGAGTGGGGGACTGTGCTCATGACAGATGAGCACAAAGGAATGCAGCT 1560  
QY 1561 GTTGTGTGTAACAAATTTGGGATGTGGAAGCCCTATGCAATGTTTGGTATGACCTATTTT 1620  
DB 1561 GTTGTGTGTAACAAATTTGGGATGTGGAAGCCCTATGCAATGTTTGGTATGACCTATTTT 1620  
QY 1621 AAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTTGGAATGAGTCAAAAT 1680  
DB 1621 AAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTCTTGCATTTGGAATGAGTCAAAAT 1680  
QY 1681 ATCTGGGACTGTGAACACAGTGGATGGGAAAGACATAATTTGTACACAGAGAGGATGTG 1740  
DB 1681 ATCTGGGACTGTGAACACAGTGGATGGGAAAGACATAATTTGTACACAGAGAGGATGTG 1740  
QY 1741 ATTTAAACCTGCTCAGGCTGATGCAACATGGGCTGAGGCTGGTGGCGGCGAGCAACCGC 1800  
DB 1741 ATTTAAACCTGCTCAGGCTGATGCAACATGGGCTGAGGCTGGTGGCGGCGAGCAACCGC 1800  
QY 1801 TCGTCGGGAGAGTGGAGGTGACTTTCAAGGACGCTGGGACACAGTGTGATGACGCG 1860  
DB 1801 TCGTCGGGAGAGTGGAGGTGACTTTCAAGGACGCTGGGACACAGTGTGATGACGCG 1860  
QY 1861 TGAACAGTAAAGCTGCAGCTGTGCTGTAGCCAGCTGGACCTCTTCTATCATTT 1920  
DB 1861 TGAACAGTAAAGCTGCAGCTGTGCTGTAGCCAGCTGGACCTCTTCTATCATTT 1920  
QY 1921 GGCATGGGCTGGGAACCGCTTCTACAGATATGGAANAATTTGGCTCGATGATTTTC 1980  
DB 1921 GGCATGGGCTGGGAACCGCTTCTACAGATATGGAANAATTTGGCTCGATGATTTTC 1980  
QY 1981 TGTGATGGAGATGAGTCAAGATCTCTGCTCATGCAAGCAAGTGGTGGGAAATAATGAC 2040  
DB 1981 TGTGATGGAGATGAGTCAAGATCTCTGCTCATGCAAGCAAGTGGTGGGAAATAATGAC 2040  
QY 2041 TGCAGTCAAGTGAAGATTTGGAGTGTCTGTTCTGATGCAATGCGATATGAGGCTGAGG 2100  
DB 2041 TGCAGTCAAGTGAAGATTTGGAGTGTCTGTTCTGATGCAATGCGATATGAGGCTGAGG 2100  
QY 2101 CTTTGGGTTGGAAGCAGCAGTGTGCTGGAANAAGTTGAGGTGATGCTCCAGGTCGCCGTG 2160  
DB 2101 CTTTGGGTTGGAAGCAGCAGTGTGCTGGAANAAGTTGAGGTGATGCTCCAGGTCGCCGTG 2160  
QY 2161 GGAATTTCTGTGCTTAATGGCTGGGAATGAACATTTGCTGAAGTGTGTCAGGCAACTT 2220  
DB 2161 GGAATTTCTGTGCTTAATGGCTGGGAATGAACATTTGCTGAAGTGTGTCAGGCAACTT 2220  
QY 2221 GAATGTGGTCTGCAATCAGGCTCCAGAGAGCCCTCAATTTTCAGAGAAACATTTACAC 2280  
DB 2221 GAATGTGGTCTGCAATCAGGCTCCAGAGAGCCCTCAATTTTCAGAGAAACATTTACAC 2280  
QY 2281 ATCTTAATGCAATTTCTGGCTGCACTGGAGGAGAGCCCTCTCTCTGGGATTTGATACGA 2340  
DB 2281 ATCTTAATGCAATTTCTGGCTGCACTGGAGGAGAGCCCTCTCTCTGGGATTTGATACGA 2340  
QY 2341 TGGAGTGGAAACAGACTGCTGCTCATTTAAATATGGAAGCAAGTTTGTCTCTCAGCC 2400





Db 1201 ACAATATGTGACCAGAACTGGAAGAATCAACAAGCCCTTGTGGTTTGTAAAGCAGCTAGGA 1260  
Qy 1261 TGTCCGTTACAGCTCTTTGGCAGTCGTCGTCTAAACCTAGTAATGAAGCTAGAGACATT 1320  
Db 1261 TGTCCGTTACAGCTCTTTGGCAGTCGTCGTCTAAACCTAGTAATGAAGCTAGAGACATT 1320  
Qy 1321 TGGATAAACAGCATATCTTGCACCTGGGAATGAGTCAGCTCTCTGGACTGCACATATGAT 1380  
Db 1321 TGGATAAACAGCATATCTTGCACCTGGGAATGAGTCAGCTCTCTGGACTGCACATATGAT 1380  
Qy 1381 GGAAGCAAGCAAGCAATGCTTCGGAAGATGAGTCAGTCGTCTGATGAAG 1440  
Db 1381 GGAAGCAAGCAAGCAATGCTTCGGAAGATGAGTCAGTCGTCTGATGAAG 1440  
Qy 1441 GCAGATCTGGACCTAAGCTTGTTCGGGCTCATAGCCCTCTTATGGAGATGAGAGTG 1500  
Db 1441 GCAGATCTGGACCTAAGCTTGTTCGGGCTCATAGCCCTCTTATGGAGATGAGAGTG 1500  
Qy 1501 AAATACCAAGGAGATGGGGACTGTGTGTCATCAGAGATGGACACAAAGGAATGCAGCT 1560  
Db 1501 AAATACCAAGGAGATGGGGACTGTGTGTCATCAGAGATGGACACAAAGGAATGCAGCT 1560  
Qy 1561 GTTCTGTGTAACAAATTTGGGATGTGGAAGCCTATGATGTGTTGGTATGACCTATTTT 1620  
Db 1561 GTTCTGTGTAACAAATTTGGGATGTGGAAGCCTATGATGTGTTGGTATGACCTATTTT 1620  
Qy 1621 AAAGAAGCATCAGGACCTATTGGCTGGATGACCTTTCTTCATTGGAATGAGTCAAT 1680  
Db 1621 AAAGAAGCATCAGGACCTATTGGCTGGATGACCTTTCTTCATTGGAATGAGTCAAT 1680  
Qy 1681 ATCTGGGACTGTGAACACAGTGGATGGGAAAGCATAAATGTGTACACAGAGAGATGTG 1740  
Db 1681 ATCTGGGACTGTGAACACAGTGGATGGGAAAGCATAAATGTGTACACAGAGAGATGTG 1740  
Qy 1741 ATTGTACCTGCTCAGTGATGCAACATGGGCCCTGAGGCTGTGGGGGAGCAACCGC 1800  
Db 1741 ATTGTACCTGCTCAGTGATGCAACATGGGCCCTGAGGCTGTGGGGGAGCAACCGC 1800  
Qy 1801 TGCTCGGGAAGACTGGAGGTGACTTTCAAGGACGGTGGGCAACAGTGTGTGATGACGGC 1860  
Db 1801 TGCTCGGGAAGACTGGAGGTGACTTTCAAGGACGGTGGGCAACAGTGTGTGATGACGGC 1860  
Qy 1861 TGGAAACAGTAAAGCTGACGTGTGGTGTAGCCAGCTGGACTGCCCATCTTCTATCAT 1920  
Db 1861 TGGAAACAGTAAAGCTGACGTGTGGTGTAGCCAGCTGGACTGCCCATCTTCTATCAT 1920  
Qy 1921 GGCATGGGCTGGGAAACGCTTCTACAGGATATGGAATAATTTGGCTCGATGTTTCC 1980  
Db 1921 GGCATGGGCTGGGAAACGCTTCTACAGGATATGGAATAATTTGGCTCGATGTTTCC 1980  
Qy 1981 TGTGATGGAGATGAGTCAGATCTCTGGTCAATGAGAACAGTGGTGGGAAATTAATGAC 2040  
Db 1981 TGTGATGGAGATGAGTCAGATCTCTGGTCAATGAGAACAGTGGTGGGAAATTAATGAC 2040  
Qy 2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTGTCATGATCGGATATGAGCTAGG 2100  
Db 2041 TGCAGTCACAGTGAAGATGTTGGAGTGATCTGTGTCATGATCGGATATGAGCTAGG 2100  
Qy 2101 CTTCTGGGTGGAAGCAGCAGTGTCGGAAGCTTCAAGGATGAGTGGGATGTCAGGGTGCCTG 2160  
Db 2101 CTTCTGGGTGGAAGCAGCAGTGTCGGAAGCTTCAAGGATGAGTGGGATGTCAGGGTGCCTG 2160  
Qy 2161 GGAATTCGTGCTGCTAATGGCTGGGAATGAACATTTGCTGAAGTTGTTGACAGCAACTT 2220  
Db 2161 GGAATTCGTGCTGCTAATGGCTGGGAATGAACATTTGCTGAAGTTGTTGACAGCAACTT 2220  
Qy 2221 GAATGTGGGCTGCAATCAGGCTTCCAGAGAGCCTCATTTTACAGAAAGAACATTACAC 2280  
Db 2221 GAATGTGGGCTGCAATCAGGCTTCCAGAGAGCCTCATTTTACAGAAAGAACATTACAC 2280  
Qy 2281 ATCTTAATGCGAATTCGTGGCTGACATGGAGGGGAAGCCTCTCTCTGGGATGTATAGGA 2340  
Db 2281 ATCTTAATGCGAATTCGTGGCTGACATGGAGGGGAAGCCTCTCTCTGGGATGTATAGGA 2340

Qy 2341 TGGAGTGGAAACAGACTGCGTGTCTATTAAATATGGAAGCAAGTTTGTATCTGCTCAGCC 2400  
Db 2341 TGGAGTGGAAACAGACTGCGTGTCTATTAAATATGGAAGCAAGTTTGTATCTGCTCAGCC 2400  
Qy 2401 CACAGGAGCCAGGCTGGTTGGAGCTGATATGCCCTCTCTGGAGCTGTTGAAGTGA 2460  
Db 2401 CACAGGAGCCAGGCTGGTTGGAGCTGATATGCCCTCTCTGGAGCTGTTGAAGTGA 2460  
Qy 2461 CATCGACACATGGCGCTCTGTCTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520  
Db 2461 CATCGACACATGGCGCTCTGTCTGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520  
Qy 2521 CTGTGACAGAAATTAATTTGGAGATGCCATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2580  
Db 2521 CTGTGACAGAAATTAATTTGGAGATGCCATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2580  
Qy 2581 AAAGGAATGCTTAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2640  
Db 2581 AAAGGAATGCTTAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2640  
Qy 2641 GCATTATGCCCATTTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700  
Db 2641 GCATTATGCCCATTTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTTGGA 2700  
Qy 2701 GTTGTCTCTCCGATATACAGATGTCGAGCTTGTGAATGGCAATCCAGTGTGACGGG 2760  
Db 2701 GTTGTCTCTCCGATATACAGATGTCGAGCTTGTGAATGGCAATCCAGTGTGACGGG 2760  
Qy 2761 CAAGTGGAGATCAACAGTCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820  
Db 2761 CAAGTGGAGATCAACAGTCTTGGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCA 2820  
Qy 2821 GAAGATGCCCCGTCTATGACAGACAGCTCAGCTGTGGAGCTGTCTCTCAACACACAGA 2880  
Db 2821 GAAGATGCCCCGTCTATGACAGACAGCTCAGCTGTGGAGCTGTCTCTCAACACACAGA 2880  
Qy 2881 CGAAATATATTTGGAGAAAGTGTCTGTGTGGGACACAGGTTTTCATTGCTTAGGG 2940  
Db 2881 CGAAATATATTTGGAGAAAGTGTCTGTGTGGGACACAGGTTTTCATTGCTTAGGG 2940  
Qy 2941 AATGAGTCACTTCTGGATAACTGTCAAAATGACAGTCTTGGAGCAGCTCCCTGTATCCAT 3000  
Db 2941 AATGAGTCACTTCTGGATAACTGTCAAAATGACAGTCTTGGAGCAGCTCCCTGTATCCAT 3000  
Qy 3001 GGAATACTGCTCTGTGATCTGCAGAGAACCTTGACCCAGCCACTGTTCATGCGCTC 3060  
Db 3001 GGAATACTGCTCTGTGATCTGCAGAGAACCTTGACCCAGCCACTGTTCATGCGCTC 3060  
Qy 3061 GCAATGTATCTGACCCCATATTTGTCTCAGTTCCAGAGGCGAGTGTTCATCTGCTTA 3120  
Db 3061 GCAATGTATCTGACCCCATATTTGTCTCAGTTCCAGAGGCGAGTGTTCATCTGCTTA 3120  
Qy 3121 CAGGACAAACGGCTCCGCTAGTGGATGGGACACCCCTGTGCCGGAGAGTAGATC 3180  
Db 3121 CAGGACAAACGGCTCCGCTAGTGGATGGGACACCCCTGTGCCGGAGAGTAGATC 3180  
Qy 3181 TATCAGCAGCGCTTCTGGGCGACCATCTGTATGACGCTGGGACCTGAGCGATGCCAC 3240  
Db 3181 TATCAGCAGCGCTTCTGGGCGACCATCTGTATGACGCTGGGACCTGAGCGATGCCAC 3240  
Qy 3241 GTGGTGTCTAAAAGCTGGGCTGTGGAGTGGCCCTTCAATGCCAGGCTCTCTGCTCATT 3300  
Db 3241 GTGGTGTCTAAAAGCTGGGCTGTGGAGTGGCCCTTCAATGCCAGGCTCTCTGCTCATT 3300  
Qy 3301 GGGAGGGGCTCAGGCGCCCTCTGGCTGGATGACCTGAACTGCACAGGAACGAGTCCAC 3360  
Db 3301 GGGAGGGGCTCAGGCGCCCTCTGGCTGGATGACCTGAACTGCACAGGAACGAGTCCAC 3360  
Qy 3361 TTGTGGCAGTGGCTTCCCGGGCTGGGGGAGCAGCAGCTGCAGGCAACAGGAGACCA 3420  
Db 3361 TTGTGGCAGTGGCTTCCCGGGCTGGGGGAGCAGCAGCTGCAGGCAACAGGAGACCA 3420



QY 508 AACAACTCCCTGTTTCAGGAGAGTGGAGTGAAATTCCAAGAAAGTGGGGACTATATGT 567  
DB 481 AACAACTCCCTGTTTCAGGAGAGTGGAGTGAAATTCCAAGAAAGTGGGGACTATATGT 540  
QY 568 GATGATGGGTGGAACATTGAATACATGCTGCCGTGGTGTGCAGGCAACATGAGATGCCATCT 627  
DB 541 GATGATGGGTGGAACATTGAATACATGCTGCCGTGGTGTGCAGGCAACATGAGATGCCATCT 600  
QY 628 TCCTTTATTTCTCTGAGTGTGTAATAGCCCTGCTATATGCCCCCAATTGGCTGGAT 687  
DB 601 TCCTTTATTTCTCTGAGTGTGTAATAGCCCTGCTATATGCCCCCAATTGGCTGGAT 660  
QY 688 GACATTTTATGCCAGGGGAATGAGTGGCACTCTGGAATTCGACACATCTGGTGGATGGGA 747  
DB 661 GACATTTTATGCCAGGGGAATGAGTGGCACTCTGGAATTCGACACATCTGGTGGATGGGA 720  
QY 748 AATCATGACTGCACTCAATAGGATGTGCACATTAATCTGTTATGATAGTAGTATCTT 807  
DB 721 AATCATGACTGCACTCAATAGGATGTGCACATTAATCTGTTATGATAGTAGTATCTT 780  
QY 808 GAACATAGGCTTGTAGTGGAACTAACCGCTGTATGGGAGAGTAGAGCTGAATATCCAA 867  
DB 781 GAACATAGGCTTGTAGTGGAACTAACCGCTGTATGGGAGAGTAGAGCTGAATATCCAA 840  
QY 868 GGAAGGTGGGGACCGTATGCCACATAAGTGGAAACAATGCTGCAGCTGATGCTGATGC 927  
DB 841 GGAAGGTGGGGACCGTATGCCACATAAGTGGAAACAATGCTGCAGCTGATGCTGATGC 900  
QY 928 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCAATTTGCACTCAGG 987  
DB 901 AAGCAGTTGGGATGTGGAACCGCACTTCACTTCGCTGGCTTGCCTCAATTTGCACTCAGG 960  
QY 988 TCTGATCTGTATGGTGTGATGGTGTCTCTGCTCCGGTATGATGATCTTTCTTTGGGAC 1047  
DB 961 TCTGATCTGTATGGTGTGATGGTGTCTCTGCTCCGGTATGATGATCTTTCTTTGGGAC 1020  
QY 1048 TGCAGACATTCGCGAACCGTCAATTTTGACTGTCTTTCATCAAAACGATGTCTGTGATC 1107  
DB 1021 TGCAGACATTCGCGAACCGTCAATTTTGACTGTCTTTCATCAAAACGATGTCTGTGATC 1080  
QY 1108 TGCTCAGATGAGCAGATTTGGAACTCGGACTACGAGATGGAAGTACAATTTGTCAGG 1167  
DB 1081 TGCTCAGATGAGCAGATTTGGAACTCGGACTACGAGATGGAAGTACAATTTGTCAGG 1140  
QY 1168 AGACTAGAGTGGAATTCATGACAGTGTGGCAATATGTCAGCAGAACTGCAAGAT 1227  
DB 1141 AGACTAGAGTGGAATTCATGACAGTGTGGCAATATGTCAGCAGAACTGCAAGAT 1200  
QY 1228 GAACAAGCCCTTGTGGTGTGTAAGCAGCTAGGATGTCCGTTTCAGCGTCTTTGGCAGTCTG 1287  
DB 1201 GAACAAGCCCTTGTGGTGTGTAAGCAGCTAGGATGTCCGTTTCAGCGTCTTTGGCAGTCTG 1260  
QY 1288 CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGACATATCTTGCAGCTGG 1347  
DB 1261 CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGACATATCTTGCAGCTGG 1320  
QY 1348 AATCAGTCAGCTCTCGGACTGCACATATGATGGAAGCAAGCAAGCAACATGCTCCGA 1407  
DB 1321 AATCAGTCAGCTCTCGGACTGCACATATGATGGAAGCAAGCAAGCAACATGCTCCGA 1380  
QY 1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGCAGATCTGGACCTTAAGCCTTGTCCGG 1467  
DB 1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAAGCAGATCTGGACCTTAAGCCTTGTCCGG 1440  
QY 1468 GCTCATAGCCCTGTTATGGGAGATTTGGAGTGAAATACCAAGGAGTGGGGAGCTGTG 1527  
DB 1441 GCTCATAGCCCTGTTATGGGAGATTTGGAGTGAAATACCAAGGAGTGGGGAGCTGTG 1500  
QY 1528 TGTATCAGATGAGCAGACAGGAATGCACTGCTGTGTGTAACAAATTTGGATGTGGA 1587  
DB 1501 TGTATCAGATGAGCAGACAGGAATGCACTGCTGTGTGTAACAAATTTGGATGTGGA 1560

QY 1588 AAGCCTATGATGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTGGCTG 1647  
DB 1561 AAGCCTATGATGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTGGCTG 1620  
QY 1648 GATGACGTTTCTTGCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1707  
DB 1621 GATGACGTTTCTTGCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1680  
QY 1708 GGAAGACATAATCTGTACACAGAGAGGATGTGATTAACCTGCTCAGGTGATGCAACA 1767  
DB 1681 GGAAGACATAATCTGTACACAGAGAGGATGTGATTAACCTGCTCAGGTGATGCAACA 1740  
QY 1768 TGGGSCCTGAGGCTGGTGGCGGCGAGCAACCGCTGCTCGGAAGACTCGAGGTGACTTT 1827  
DB 1741 TGGGSCCTGAGGCTGGTGGCGGCGAGCAACCGCTGCTCGGAAGACTCGAGGTGACTTT 1800  
QY 1828 CAAGGACGTTGGGCGACAGTGTGTATGACGCTGGAACAGTAAAGCTGCAAGCTGTGGT 1887  
DB 1801 CAAGGACGTTGGGCGACAGTGTGTATGACGCTGGAACAGTAAAGCTGCAAGCTGTGGT 1860  
QY 1888 TGTAGCCAGCTGGAGTGGCCCATCTTCTATCATTTGGCATGGGTCTGGAAACGCTTCTACA 1947  
DB 1861 TGTAGCCAGCTGGAGTGGCCCATCTTCTATCATTTGGCATGGGTCTGGAAACGCTTCTACA 1920  
QY 1948 GGATATGGAATAATTTGGCTCGATGATGTTTCTGTTGATGAGATGAGTCAAGTCTCTGG 2007  
DB 1921 GGATATGGAATAATTTGGCTCGATGATGTTTCTGTTGATGAGATGAGTCAAGTCTCTGG 1980  
QY 2008 TCATGACGGAACAGTGGTGGGGAATAATGACTGCAGTCAAGTGAAGATGTTGGAGT 2067  
DB 1981 TCATGACGGAACAGTGGTGGGGAATAATGACTGCAGTCAAGTGAAGATGTTGGAGT 2040  
QY 2068 ATCTGTTCTGATGATCGATGATGAGCTGAGGCTTGTGGTGGAAACGAGTGTGCT 2127  
DB 2041 ATCTGTTCTGATGATCGATGATGAGCTGAGGCTTGTGGTGGAAACGAGTGTGCT 2100  
QY 2128 GGAAGAAGTTGAGTGAATGTCCAGGTCGCTGGGAATTCGTGTGCTTAATGGCTGGGA 2187  
DB 2101 GGAAGAAGTTGAGTGAATGTCCAGGTCGCTGGGAATTCGTGTGCTTAATGGCTGGGA 2160  
QY 2188 ATGAACATTTGCTGAAGTGTGTTTGCAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCC 2247  
DB 2161 ATGAACATTTGCTGAAGTGTGTTTGCAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCC 2220  
QY 2248 AGAGAGCCTCATTTTCACAGAAAGAACATTACACATCTTAATGCTCGAATTTCTGGCTGCACT 2307  
DB 2221 AGAGAGCCTCATTTTCACAGAAAGAACATTACACATCTTAATGCTCGAATTTCTGGCTGCACT 2280  
QY 2308 GGAGGGAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTGCGTGTCTAT 2367  
DB 2281 GGAGGGAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTGCGTGTCTAT 2340  
QY 2368 TTAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCAGCAGCCAGGCTGGTGGAGCT 2427  
DB 2341 TTAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCAGCAGCCAGGCTGGTGGAGCT 2400  
QY 2428 GATATGCCCTGCTCTGGACGTGTTTGAAGTGAACATGACAGACATGCGCTCTCTGCTGT 2487  
DB 2401 GATATGCCCTGCTCTGGACGTGTTTGAAGTGAACATGACAGACATGCGCTCTCTGCTGT 2460  
QY 2488 GATTCTGATTTCTCTCTATGCTGCCAATGTGCTGTGTCAGAGAAATTAATTTGTTGGAGAT 2547  
DB 2461 GATTCTGATTTCTCTCTATGCTGCCAATGTGCTGTGTCAGAGAAATTAATTTGTTGGAGAT 2520  
QY 2548 GCCATATCTTCTCTGTTGGGAGATCACTTTGGAAGGGAATGCTTAACCTTGGGCCGAA 2607  
DB 2521 GCCATATCTTCTCTGTTGGGAGATCACTTTGGAAGGGAATGCTTAACCTTGGGCCGAA 2580  
QY 2608 AAGTCCAGTGTGAAGGAGTGAACCTACCTTTGATTTATGCCCATTTGTTCAACATCCG 2667  
DB 2581 AAGTCCAGTGTGAAGGAGTGAACCTACCTTTGATTTATGCCCATTTGTTCAACATCCG 2640  
QY 2668 GAAGACACTTGTATCCACAGCAGAGAAAGTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727

Db 2641 GAAGACACTGTGATCCACAGCAGAGAGTTGGAGTTGCTGTTCCGATATACAGATGTC 2700  
QY 2728 CGACTTGTGAATGGCAAAATCCAGTGTGACGGGCAAGTGGAGATCAACGCTGCTGGACAC 2787  
Db 2701 CGACTTGTGANTGGCAAAATCCAGTGTGACGGGCAAGTGGAGATCAACGCTGCTGGACAC 2760  
QY 2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATGCCGCTGTTCTATGACAGAC 2847  
Db 2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATGCCGCTGTTCTATGACAGAC 2820  
QY 2848 CTCAGCTGTGGGACTGCTCTCAACCCACAGGAGGAAATATATGGAGAAAGAGTGT 2907  
Db 2821 CTCAGCTGTGGGACTGCTCTCAACCCACAGGAGGAAATATATGGAGAAAGAGTGT 2880  
QY 2908 CGTGTGTGGGACACAGGTTTCATCTCTTAGGGAATGAGTCACTTCTGGATACTGTCAA 2967  
Db 2881 CGTGTGTGGGACACAGGTTTCATCTCTTAGGGAATGAGTCACTTCTGGATACTGTCAA 2940  
QY 2968 ATGACAGTCTTTGGAGCACCCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3027  
Db 2941 ATGACAGTCTTTGGAGCACCCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCTGCACA 3000  
QY 3028 GGNAGCCTGACCCAGCAGCTGTTTCCATGCTCGGCAAAATGTATCTGACCCATATTTGCT 3087  
Db 3001 GGAAGCCTGACCCAGCAGCTGTTTCCATGCTCGGCAAAATGTATCTGACCCATATTTGCT 3060  
QY 3088 GCAGTTTCCAGAGGGCAGTGTGATCTGCTTAGAGGACAAACGGCTCCGCTCTAGTGGAT 3147  
Db 3061 GCAGTTTCCAGAGGGCAGTGTGATCTGCTTAGAGGACAAACGGCTCCGCTCTAGTGGAT 3120  
QY 3148 GGGGACAGCGCTGTGCCGGGAGAGTAGATCTATCAGACGGCTTCTGGGGCACCATC 3207  
Db 3121 GGGGACAGCGCTGTGCCGGGAGAGTAGATCTATCAGACGGCTTCTGGGGCACCATC 3180  
QY 3208 TGTGATGAGCGCTGGGACCTGAGCGATGCCAGTGGTGTCAAAAGCTGGCTGTGA 3267  
Db 3181 TGTGATGAGCGCTGGGACCTGAGCGATGCCAGTGGTGTCAAAAGCTGGCTGTGA 3240  
QY 3268 GTGGCCTTCAATGCCAGGCTCTGCTCACTTTGGGGAGGGGTGAGGGCCCACTCTGGCTG 3327  
Db 3241 GTGGCCTTCAATGCCAGGCTCTGCTCACTTTGGGGAGGGGTGAGGGCCCACTCTGGCTG 3300  
QY 3328 GATGACCTGAATGACAGAGACGGAGTCCACTTGTGGAGTGCCCTTCCCGCGCTGG 3387  
Db 3301 GATGACCTGAATGACAGAGACGGAGTCCCACTTGTGGCAGTGCCTTCCCGCGCTGG 3360  
QY 3388 GGGCAGCAGCTGCGAGGACAGGAGGACGAGGGTCTATCTGCTCAGAAATTCAGACC 3447  
Db 3361 GGGCAGCAGCTGCGAGGACAGGAGGACGAGGGTCTATCTGCTCAGAAATTCAGACC 3420  
QY 3448 TTGAGGCTCTACAGTGAATGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507  
Db 3421 TTGAGGCTCTACAGTGAATGAAACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480  
QY 3508 AACGGACCTGGGGCAGCTGCGCAGAGAGACATCACCACAGCCATACAGGCAATGTG 3567  
Db 3481 AACGGACCTGGGGCAGCTGCGCAGAGAGAGATCACCACAGCCATACAGGCAATGTG 3540  
QY 3568 TGCAGCAGCTGGGCTGTGGGAGAGTGGAGTTGTCAGCTCGCCCTTATCTAAGACA 3627  
Db 3541 TGCAGCAGCTGGGCTGTGGGAGAGTGGAGTTGTCAGCTCGCCCTTATCTAAGACA 3600  
QY 3628 GGCTCTGTTTCAATGTTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATATGG 3687  
Db 3601 GGCTCTGTTTCAATGTTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATATGG 3660  
QY 3688 CAGTGGCTGTCTCCCGATGGGAGGAGAAATCTCAGCCAGCAGAGAGACCTGGATC 3747  
Db 3661 CAGTGGCTGTCTCCCGATGGGAGGAGAAATCTCAGCCAGCAGAGAGACCTGGATC 3720  
QY 3748 ACATGTGAGATAGATAGAGTGGCTGGAGGAGACACGAGTGTCTGGAGAGTGGAG 3807

Db 3721 ACATGTGAGATAGATAGATAGAGTGGCTGGAGGAGACACGAGTGTCTGGAGAGTGGAG 3780  
QY 3808 ATCTGSCACGACAGCTCTGGGGACACAGTGTGTGATGACTCTCTGGACCTGGCCGAGGCG 3867  
Db 3781 ATCTGSCACGACAGCTCTGGGGACACAGTGTGTGATGACTCTCTGGACCTGGCCGAGGCG 3840  
QY 3868 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTGTCTCTGGCTGTGGCTGTGGCTGTGGCTGTGG 3927  
Db 3841 GAAGTGGTGTGTCAGCAGCTGGGCTGTGGCTGTCTCTGGCTGTGGCTGTGGCTGTGGCTGTGG 3900  
QY 3928 TTTGGCCAGGGAACTGGAACCATCTGTTGGATGACATCGGTGCAAAAGAAATGAGTCA 3987  
Db 3901 TTTGGCCAGGGAACTGGAACCATCTGTTGGATGACATCGGTGCAAAAGAAATGAGTCA 3960  
QY 3988 TTTCTATGGACTGTCTCACGCCAAACCTGGGACACAGTGTGTGACACAAAGGAAGAT 4047  
Db 3961 TTTCTATGGACTGTCTCACGCCAAACCTGGGACACAGTGTGTGACACAAAGGAAGAT 4020  
QY 4048 GCTGGGCTGAGTGTCTGTGACAGTGTGCTGAAATCACTGAATGCTCTCAGTCAATTTA 4107  
Db 4021 GCTGGGCTGAGTGTCTGTGACAGTGTGCTGAAATCACTGAATGCTCTCAGTCAATTTA 4080  
QY 4108 GCAGTTATTTATCCAGTATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4167  
Db 4081 GCAGTTATTTATCCAGTATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4140  
QY 4168 AGTGTGCTGGAGTGTCTGAGAAACAAACAACTCTGCCCTCAGAGTTTCAACCAAGAGGAGG 4227  
Db 4141 AGTGTGCTGGAGTGTCTGAGAAACAAACAACTCTGCCCTCAGAGTTTCAACCAAGAGGAGG 4200  
QY 4228 GGTCTCTCAGGAGAAATTTATTCATGAGATGGAGACCTGCTCAAGAGAGGACCCA 4287  
Db 4201 GGTCTCTCAGGAGAAATTTATTCATGAGATGGAGACCTGCTCAAGAGAGGACCCA 4260  
QY 4288 CATGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4347  
Db 4261 CATGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4320  
QY 4348 TCGCTCTGGAGTGTCT 4386  
Db 4321 TCGCTCTGGAGTGTCT 4359

RESULT 4

US-09-759-130B-380  
; Sequence 380, Application US/09759130B  
; Publication No. US2003002279A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: McCarthy, Sean A  
; APPLICANT: Fraser, Christopher C  
; APPLICANT: Sharp, John D  
; APPLICANT: Barnes, Thomas S  
; APPLICANT: Kirst, Susan J  
; APPLICANT: Mackay, Charles R  
; APPLICANT: Myers, Paul S  
; APPLICANT: Leiby, Kevin R  
; APPLICANT: Wrighton, Nicolas  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Holtzman, Douglas A  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER  
; FILE REFERENCE: MPI00-5350MNI  
; CURRENT APPLICATION NUMBER: US/09/759,130B  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: US 09/479,249  
; PRIOR FILING DATE: 2000-01-07  
; PRIOR APPLICATION NUMBER: US 09/559,497  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: US 09/333,159

```

: PRIOR FILING DATE: 1999-06-14
: PRIOR APPLICATION NUMBER: US 09/596,194
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 09/342,364
: PRIOR FILING DATE: 1999-06-29
: PRIOR APPLICATION NUMBER: US 09/608,452
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/393,996
: PRIOR FILING DATE: 1999-09-10
: PRIOR APPLICATION NUMBER: US 09/602,871
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: US 09/420,707
: PRIOR FILING DATE: 1999-10-19
: NUMBER OF SEQ ID NOS: 460
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 380
: LENGTH: 4359
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-759-1308-380

```

Query Match	94.2%	Score 4359;	DB 9;	Length 4359;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4359;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	28	ATGATGCTGCTCAAAACTCGTGGCATATTGATTTTGGAAAGATGCTGCTGTCATCAGAAC	87	
DB	1	ATGATGCTGCTCAAAACTCGTGGCATATTGATTTTGGAAAGATGCTGCTGTCATCAGAAC	60	
QY	88	CTTTTCTCTGCTGGGTAACTTGCATCCTGCTCCTGATTCCTGCTTCTTCATCAGAGT	147	
DB	61	CTTTTCTCTGCTGGGTAACTTGCATCCTGCTCCTGATTCCTGCTTCTTCATCAGAGT	120	
QY	148	TTTAATGGAAACAGATTGGAGTTTCAGGCTGGTCAATGGAGACGCTCCCTGCTCTGGGACA	207	
DB	121	TTTAATGGAAACAGATTGGAGTTTCAGGCTGGTCAATGGAGACGCTCCCTGCTCTGGGACA	180	
QY	208	GTGGAGGTGAAATTCACAGGACACAGTGGGGACACTGCTGTGATGATGGGTGAAACACTACT	267	
DB	181	GTGGAGGTGAAATTCACAGGACACAGTGGGGACACTGCTGTGATGATGGGTGAAACACTACT	240	
QY	268	GCCTCAACTGCTGCTGGTGGAAACACGCTTGGATGCCATTTCTTTCGGCATGTTTCGTTTT	327	
DB	241	GCCTCAACTGCTGCTGGTGGAAACACGCTTGGATGCCATTTCTTTCGGCATGTTTCGTTTT	300	
QY	328	GGACAAGCCGTGACTAGACATGGAANAATTTGGCTTGTGATGATGTTTCTGTATTGGAAT	387	
DB	301	GGACAAGCCGTGACTAGACATGGAANAATTTGGCTTGTGATGATGTTTCTGTATTGGAAT	360	
QY	388	GAGTCAGCTCTCTGGGAATGTCACACCGGAATGGGGAAGCCCAACTGTTATCATGGA	447	
DB	361	GAGTCAGCTCTCTGGGAATGTCACACCGGAATGGGGAAGCCCAACTGTTATCATGGA	420	
QY	448	GAAGATGTTGGTGTGAACCTGTTATGGTGAAGCCCAACTGGGTTTGGAGCTAGTGGATGGA	507	
DB	421	GAAGATGTTGGTGTGAACCTGTTATGGTGAAGCCCAACTGGGTTTGGAGCTAGTGGATGGA	480	
QY	508	AACAACCTCCTGTTCCAGGGAGAGTGGAGGTGAAATTTCCAAGAAAGGTGGGGACATATATGT	567	
DB	481	AACAACCTCCTGTTCCAGGGAGAGTGGAGGTGAAATTTCCAAGAAAGGTGGGGACATATATGT	540	
QY	568	GATGATGGGTGGAACTTGAATFACTGCTGCCGTGCTGTGTCAGGCAACTAGGATGTCCATCT	627	
DB	541	GATGATGGGTGGAACTTGAATFACTGCTGCCGTGCTGTGTCAGGCAACTAGGATGTCCATCT	600	
QY	628	TCTTTTATTCTCTGGAGTTGTTAAATAGCCCTGCTGTATTGCCCCCATTTGGCTGGAT	687	
DB	601	TCTTTTATTCTCTGGAGTTGTTAAATAGCCCTGCTGTATTGCCCCCATTTGGCTGGAT	660	
QY	688	GACATTTTATCCCAAGGGAATGAGTTGGCACTCTGGAAATTCACACATCTGGATGGGA	747	
DB	661	GACATTTTATGCCAGGGAATGAGTTGGCACTCTGGAAATTCACACATCTGGATGGGA	720	

Qy	748	AATCATGACGTGCAGTGCACAAATGAGGAGTGCACATTAACCTGTTATGATAGTAGTGATCTTT	807
Db	749		
Qy	750		
Db	751	AATCATGACGTGCAGTGCACAAATGAGGAGTGCACATTAACCTGTTATGATAGTAGTGATCTTT	780
Qy	752		
Qy	808	GAACTAAGGCTTCTAGCTGGAACTAACCGCTGTATGGGAGAGTAGAGCTGAAATCCAA	867
Db	809		
Db	781	GAACTAAGGCTTCTAGCTGGAACTAACCGCTGTATGGGAGAGTAGAGCTGAAATCCAA	840
Qy	868	GGAGGTTGGGGACCGTATGCCACCATAAGTGGAAACAATGCTGCAGCTGATGTCGTATGC	927
Db	869		
Db	841	GGAGGTTGGGGACCGTATGCCACCATAAGTGGAAACAATGCTGCAGCTGATGTCGTATGC	900
Qy	928	AAGCAGTTGGGATGCTGGAACCGCAGCTTCACTTCGCTGCGCTTGCCTCATTTGCCAGTCAGGG	987
Db	929		
Db	901	AAGCAGTTGGGATGCTGGAACCGCAGCTTCACTTCGCTGCGCTTGCCTCATTTGCCAGTCAGGG	960
Qy	988	TCGTGATGTTGATGGCTTGATGGTGTCTCCTGCTCCGGTAAATGAATCTTTCTTTGGGCAC	1047
Db	989		
Db	961	TCGTGATGTTGATGGCTTGATGGTGTCTCCTGCTCCGGTAAATGAATCTTTCTTTGGGCAC	1020
Qy	1048	TGCAGACATTTCCGGAACCGTCAATTTTGACTGCTTTCATCAAAAACGATGTGTCGTGATC	1107
Db	1049		
Db	1021	TGCAGACATTTCCGGAACCGTCAATTTTGACTGCTTTCATCAAAAACGATGTGTCGTGATC	1080
Qy	1108	TGCTCAGATGGACAGATTTGGAACTGCGACTAGCAGATAGCAAGTGAAGTAACAATTTGTTACAGG	1167
Db	1109		
Db	1081	TGCTCAGATGGACAGATTTGGAACTGCGACTAGCAGATAGCAAGTGAAGTAACAATTTGTTACAGG	1140
Qy	1168	AGAGTAGAGGTGAGAAATCATGAACAGTGGTGGCAACAATATGTGACCAAGAACTGGAAGAAT	1227
Db	1169		
Db	1141	AGAGTAGAGGTGAGAAATCATGAACAGTGGTGGCAACAATATGTGACCAAGAACTGGAAGAAT	1200
Qy	1228	GAACAAGCCCTTGCTGTTGTAAGCAGCTAGAGATGTCGGCTTCAGCGCTTTTGGCAGTCGT	1287
Db	1229		
Db	1201	GAACAAGCCCTTGCTGTTGTAAGCAGCTAGAGATGTCGGCTTCAGCGCTTTTGGCAGTCGT	1260
Qy	1288	CGTGCTAAACCTTAGTAATGAAGCTAGAGACATTTGGATTAACAGCATATCTTTGCACATGCG	1347
Db	1289		
Db	1261	CGTGCTAAACCTTAGTAATGAAGCTAGAGACATTTGGATTAACAGCATATCTTTGCACATGCG	1320
Qy	1348	AATGAGTCAGCTCTCTGGGACTGCAATATGATGAAAGCAAGCAAGCAACATGCTCCGA	1407
Db	1349		
Db	1321	AATGAGTCAGCTCTCTGGGACTGCAATATGATGAAAGCAAGCAAGCAACATGCTCCGA	1380
Qy	1408	AGATCAGATCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGCCTTTGTCGGG	1467
Db	1409		
Db	1381	AGATCAGATCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGCCTTTGTCGGG	1440
Qy	1468	GCTCATAGCCCTTGTATGGGAGATTGGAGGTGAATATCCAAAGAGAGATGGGGGACATGTG	1527
Db	1469		
Db	1441	GCTCATAGCCCTTGTATGGGAGATTGGAGGTGAATATCCAAAGAGAGATGGGGGACATGTG	1500
Qy	1528	TGTCATCAGAGATGGAGCAACAGGAATGCAGCTGTGTGCTGTAACAATTCGGATGGA	1587
Db	1529		
Db	1501	TGTCATCAGAGATGGAGCAACAGGAATGCAGCTGTGTGCTGTAACAATTCGGATGGA	1560
Qy	1588	AAGCCTATGCATGTGTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTTGGCTG	1647
Db	1589		
Db	1561	AAGCCTATGCATGTGTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTTGGCTG	1620
Qy	1648	GATGACGTTTTCTTGCATTTGSAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG	1707
Db	1649		
Db	1621	GATGACGTTTTCTTGCATTTGSAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG	1680
Qy	1708	GGAAGACATTAATTTGTTACACAGAGAGGATGTATTGTAACCTGCTCAGGTGATGCAACA	1767
Db	1709		
Db	1681	GGAAGACATTAATTTGTTACACAGAGAGGATGTATTGTAACCTGCTCAGGTGATGCAACA	1740
Qy	1768	TGGGGCTGAGGCTGTGGCGCGCACACCCCTGCTCGGGAAGACTGAGGCTGTACTTTT	1827
Db	1769		
Db	1741	TGGGGCTGAGGCTGTGGCGCGCACACCCCTGCTCGGGAAGACTGAGGCTGTACTTTT	1800
Qy	1828	CAAGGACCGTGGGGCACAGTGTGTATGACGCTGGAACAGTAAAGCTGCAGCTGTGGT	1887

|||||  
Db 1801 CRAAGGCGTGGGGCACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTCAGCGTGTGGTG 1860  
|||||  
QY 1888 TGTAGCAGCTGACGACGCGCCATCTCTATCATTTGGCATGGGTCTGGAAACGCTTCTACA 1947  
|||||  
Db 1861 TGTAGCAGCTGACGACGCGCCATCTCTATCATTTGGCATGGGTCTGGAAACGCTTCTACA 1920  
|||||  
QY 1948 GGATATGGAAAAATTTGGCTCGCATGATGTTTCCCTGTGATGGAGATGAGTCAGATCTCTGG 2007  
|||||  
Db 1921 GGATATGGAAAAATTTGGCTCGCATGATGTTTCCCTGTGATGGAGATGAGTCAGATCTCTGG 1980  
|||||  
QY 2008 TCATGACGAGAACAGTGGGTGGGGAATATATGACTGCAGTCAAGTGAAGATGTTGAGTG 2067  
|||||  
Db 1981 TCATGACGAGAACAGTGGGTGGGGAATATATGACTGCAGTCAAGTGAAGATGTTGAGTG 2040  
|||||  
QY 2068 ATCTGTCTCTGATCATCGCATGGATATGGAGCTGAGGCTTGTGGGTGGAACAGCAGAGGTGTCT 2127  
|||||  
Db 2041 ATCTGTCTCTGATCATCGCATGGATATGGAGCTGAGGCTTGTGGGTGGAACAGCAGAGGTGTCT 2100  
|||||  
QY 2128 GGAAGAGTTGAGGTGAATGTCCAGGCTGCCGTGGGAATTTCTGTGTCTTAATGCTGGGGA 2187  
|||||  
Db 2101 GGAAGAGTTGAGGTGAATGTCCAGGCTGCCGTGGGAATTTCTGTGTCTTAATGCTGGGGA 2160  
|||||  
QY 2188 ATCAACATTTGCTCAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTCAATCAGGCTCTCC 2247  
|||||  
Db 2161 ATCAACATTTGCTCAAGTTGTTTGCAGGCAACTTGAATGTGGGTCTCAATCAGGCTCTCC 2220  
|||||  
QY 2248 AGAGAGCTCATTTACAGAAAGAACATTACACATCTTAATGTCGAATTTCTGTGCTCACT 2307  
|||||  
Db 2221 AGAGAGCTCATTTACAGAAAGAACATTACACATCTTAATGTCGAATTTCTGTGCTCACT 2280  
|||||  
QY 2308 GGAGGGAAGCCCTCTCTCTGGGATTTATACGATGGAGTGGAAACAGACTCGCTGTCTAT 2367  
|||||  
Db 2281 GGAGGGAAGCCCTCTCTCTGGGATTTATACGATGGAGTGGAAACAGACTCGCTGTCTAT 2340  
|||||  
QY 2368 TTAATATGAAGCAAGTTGATCTCTCAGCCACAGCAGCCAGCGCTGTTGAGCT 2427  
|||||  
Db 2341 TTAATATGAAGCAAGTTGATCTCTCAGCCACAGCAGCCAGCGCTGTTGAGCT 2400  
|||||  
QY 2428 GATATGCCCTGCTCTGGAGTGTGAAGTGAACATGCAGACATGGGCTCTGTCTGT 2487  
|||||  
Db 2401 GATATGCCCTGCTCTGGAGTGTGAAGTGAACATGCAGACATGGGCTCTGTCTGT 2460  
|||||  
QY 2488 GATCTGATTTCTCTCTCATGCTGCAATGTCTGTGAGAGAAATTAATTTGAGAT 2547  
|||||  
Db 2461 GATCTGATTTCTCTCTCATGCTGCAATGTCTGTGAGAGAAATTAATTTGAGAT 2520  
|||||  
QY 2548 GCCATATCTTTCTCTGGGAGATCACTTTGGAAAGGGAATGCTTAACCTTGGGCCGAA 2607  
|||||  
Db 2521 GCCATATCTTTCTCTGGGAGATCACTTTGGAAAGGGAATGCTTAACCTTGGGCCGAA 2580  
|||||  
QY 2608 AAGTTCAGTGTGAAGGGAGTGAACCTCACTTTCATTTATGCCCATTTGTTCAACATCCG 2667  
|||||  
Db 2581 AAGTTCAGTGTGAAGGGAGTGAACCTCACTTTCATTTATGCCCATTTGTTCAACATCCG 2640  
|||||  
QY 2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCGCCGATATACAGATGC 2727  
|||||  
Db 2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCGCCGATATACAGATGC 2700  
|||||  
QY 2728 CGACTTGTGAATGGAATCCCACTGTGAGGGCAAGTGGAGATCAACGTTGTCACAC 2787  
|||||  
Db 2701 CGACTTGTGAATGGAATCCCACTGTGAGGGCAAGTGGAGATCAACGTTGTCACAC 2760  
|||||  
QY 2788 TGGGGTCACTGTGTGACACCCCACTGGGACCAAGAGATGCCGCTGTTCTATCGACACAG 2847  
|||||  
Db 2761 TGGGGTCACTGTGTGACACCCCACTGGGACCAAGAGATGCCGCTGTTCTATCGACACAG 2820  
|||||  
QY 2848 CTCAGCTGTGGGACTGCTCTCTCAACCCACAGGAGGAAATATATTGGAGAAAGTGT 2907  
|||||  
Db 2821 CTCAGCTGTGGGACTGCTCTCTCAACCCACAGGAGGAAATATATTGGAGAAAGTGT 2880  
|||||  
QY 2908 CGTGTGGGGACACAGTTTCACTGCTAGGGAATGAGTCACTTCTGGATACTCTCAA 2967  
|||||

Db 2881 CGTGTGTGGGGACACAGGTTTCATTCCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2940  
QY 2968 ATGACAGTTCTTTGGAGCACTCCCTGTATCCATGGAATACTGTCTCTGTGATCTGCACA 3027  
|||||  
Db 2941 ATGACAGTTCTTTGGAGCACTCCCTGTATCCATGGAATACTGTCTCTGTGATCTGCACA 3000  
|||||  
QY 3028 GGAAGCCTGACCCAGCAGCACTGTTTCCATGCCCTGGAAATGATATCTGACCCATATTTGTCT 3087  
|||||  
Db 3001 GGAAGCCTGACCCAGCAGCACTGTTTCCATGCCCTGGAAATGATATCTGACCCATATTTGTCT 3060  
|||||  
QY 3088 CGAGTTCAGAGGGCAGTCTTTGATCTCTTAGAGCAAAACGGCTCCGCTAGTGGAT 3147  
|||||  
Db 3061 CGAGTTCAGAGGGCAGTCTTTGATCTCTTAGAGCAAAACGGCTCCGCTAGTGGAT 3120  
|||||  
QY 3148 GGGGACAGCGCTGTGCCGGGAGATCTATCAGACAGGCTTCTGGGACACCATC 3207  
|||||  
Db 3121 GGGGACAGCGCTGTGCCGGGAGATCTATCAGACAGGCTTCTGGGACACCATC 3180  
|||||  
QY 3208 TGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAAAAGCTGGGCTGTGA 3267  
|||||  
Db 3181 TGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGTGTCAAAAGCTGGGCTGTGA 3240  
|||||  
QY 3268 GTGGCCTTCAATGCCACGCTCTCTGCTCACTTTGGGAGGGGTACGGCCCATCTGGCTG 3327  
|||||  
Db 3241 GTGGCCTTCAATGCCACGCTCTCTGCTCACTTTGGGAGGGGTACGGCCCATCTGGCTG 3300  
|||||  
QY 3328 GATGACCTGAACCTGCACAGAACGGAGTCCCACTTTGTGCGAGTGCCTTTCCCGCGCTGG 3387  
|||||  
Db 3301 GATGACCTGAACCTGCACAGAACGGAGTCCCACTTTGTGCGAGTGCCTTTCCCGCGCTGG 3360  
|||||  
QY 3388 GGGCAGCAGCTGCAGGCGACAGGAGCGCAGGGGTCACTGTGTCAGAAATTCAGACC 3447  
|||||  
Db 3361 GGGCAGCAGCTGCAGGCGACAGGAGCGCAGGGGTCACTGTGTCAGAAATTCAGACC 3420  
|||||  
QY 3448 TTGAGCTCTACAGTGAACCTGAACACAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3507  
|||||  
Db 3421 TTGAGCTCTACAGTGAACCTGAACACAGAGCTGTGCTGGGAGATTGGAAGTCTTCTAT 3480  
|||||  
QY 3508 AACGGACCTGGGGCAGCTGCCAGGAGAACATCACCACAGCCATAGCAGCATTTGTG 3567  
|||||  
Db 3481 AACGGACCTGGGGCAGCTGCCAGGAGAACATCACCACAGCCATAGCAGCATTTGTG 3540  
|||||  
QY 3568 TGCAGCAGCTGGGCTGTGGGAGAAATGAGTTGTGAGCTCCGCTCCGCTTATCTAGACA 3627  
|||||  
Db 3541 TGAGCAGCTGGGCTGTGGGAGAAATGAGTTGTGAGCTCCGCTCCGCTTATCTAGACA 3600  
|||||  
QY 3628 GGCTCTGGTTTCACTGTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATATGG 3687  
|||||  
Db 3601 GGCTCTGGTTTCACTGTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATATGG 3660  
|||||  
QY 3688 CAGTGCCTGTCTGCCCATGGGAGCAAGAAATCTCCAGCCCAAGAGAGACTGGATC 3747  
|||||  
Db 3661 CAGTGCCTGTCTGCCCATGGGAGCAAGAAATCTCCAGCCCAAGAGAGACTGGATC 3720  
|||||  
QY 3748 ACATGTGAGATAGATAAGAGTCCGTGGAGAGACACCCAGTCTCTGGAGAGTGGAG 3807  
|||||  
Db 3721 ACATGTGAGATAGATAAGAGTCCGTGGAGAGACACCCAGTCTCTGGAGAGTGGAG 3780  
|||||  
QY 3808 ATCTGACAGCAGGCTCTGGGGCACAGTGTGTGATGACTCTCTGGACCTGGCGGAGCG 3867  
|||||  
Db 3781 ATCTGACAGCAGGCTCTGGGGCACAGTGTGTGATGACTCTCTGGACCTGGCGGAGCG 3840  
|||||  
QY 3868 GAAGTGTGTGTGACAGCTGGGCTGTGCTCTGTGCTCTGGCTGGCTGGAGGAGCTTCG 3927  
|||||  
Db 3841 GAAGTGTGTGTGACAGCTGGGCTGTGCTCTGTGCTCTGGCTGGCTGGAGGAGCTTCG 3900  
|||||  
QY 3928 TTTGGCCAGGAACTGGAACCATCTGGTTGGATGACATCGGGTGCAGAGAAATGAGTCA 3987  
|||||  
Db 3901 TTTGGCCAGGAACTGGAACCATCTGGTTGGATGACATCGGGTGCAGAGAAATGAGTCA 3960  
|||||  
QY 3988 TTTCTATGGGACTGTGACGCCAAACCTGGGGACAGAGTGTGTCGACACAAAGGAGAT 4047  
|||||  
Db 3961 TTTCTATGGGACTGTGACGCCAAACCTGGGGACAGAGTGTGTCGACACAAAGGAGAT 4020  
|||||

QY 4048 GCTGGCGTGGAGTCTCTGGACAGTCTGTAATCACTGAATGCTCTCTCAGGTCATTTA 4107  
|||||  
Db 4021 GCTGGCGTGGAGTCTCTGGACAGTCTGTAATCACTGAATGCTCTCTCAGGTCATTTA 4080  
|||||  
QY 4108 GCACCTATTATTCAGCATATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4167  
|||||  
Db 4081 GCACCTATTATTCAGCATATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4140  
|||||  
QY 4168 ACCTGGTGGCGAGTTCAGAAACAAACATCTGCGCTCTCAGAGTTTCAACCCAGAGGAGG 4227  
|||||  
Db 4141 ACCTGGTGGCGAGTTCAGAAACAAACATCTGCGCTCTCAGAGTTTCAACCCAGAGGAGG 4200  
|||||  
QY 4228 GGTTCCTCTCAGAGGAGATTTATTCATCAGATGGAGACCTGCTCAAGAGAGAGACCCA 4287  
|||||  
Db 4201 GGTTCCTCTCAGAGGAGATTTATTCATCAGATGGAGACCTGCTCAAGAGAGAGACCCA 4260  
|||||  
QY 4288 CATGGGACAAGAACTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4347  
|||||  
Db 4261 CATGGGACAAGAACTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4320  
|||||  
QY 4348 TCGCTGTTGGAGTCT 4386  
|||||  
Db 4321 TCGCTGTTGGAGTCT 4359  
|||||  
RESULT 5  
US-10-042-431-79  
; Sequence 79, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: FRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042,431  
; PRIOR FILING DATE: 2001-10-25  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 79  
; LENGTH: 4308  
; TYPE: DNA  
; ORGANISM: Bos sp.  
US-10-042-431-79

Query Match 19.28; Score 886.8; DB 9; Length 4308;  
Best Local Similarity 54.98; Pred. No. 5.5e-233;  
Matches 2061; Conservative 0; Mismatches 1527; Indels 168; Gaps 9;

QY 483 TCTGGGTTTGAAGTGTGATGAAACAACTCTCTCTCAGGAGAGTGGAGTGAATTT 542  
|||||  
Db 75 TCTGGAGCTGAGTTGAAGATGAGTCCATCGCTGTGAGGAGAGTGGAGTGAACCA 134  
|||||  
QY 543 CCAAGAAAGTGGGGACTATATGTGATGATGGTGGAACTTGAATCTGCTGCGGTGGT 602  
|||||  
Db 135 CCAAGAGAAATGGGCAAGTGGATGTTACAGTGGACATGAAGGATGATCTGTAGT 194  
|||||  
QY 603 GTGAGCAACTAGATGCTCATCTTTTATTTCTCTGAGTGTGTATAGCCCTGC 662  
|||||  
Db 195 GTGAGCAAGCTGGGTGTGAGTGGCATTTGTTTCTGAGGAGGCTTATTTGGGCC 254  
|||||  
QY 663 TGTATTGCGCCCTTGGCTGGATGACATTTTATGCGAGGGAATGAGTTGGCACTCTG 722  
|||||  
Db 255 AGGACTGGCCCAATTTGGCTTTTGTATCTATCTATGTAAGGAGACAGTCAACTGTGAC 314  
|||||  
QY 723 GAATTCAGACATCGTGGATGGGAAATCATGACTGCAGTGCACAAATGAGGATGTCACT 782  
|||||

Db 315 TGACTGTGAGCATTTCTAATAATTAAGAGACTATCGTAATGCTATATCATGTGTCGGGA 374  
|||||  
QY 783 AACTTTGTTATGATAGTAGTATCTTGAACAAAGCTTCTAGTGGAACTAAACCGCTGAT 842  
|||||  
Db 375 TGCTGGAGTAGTCTGCTCAGGATTG---TCCGTCTGGCTGGAGGAGTGAACCCCTGCTC 431  
|||||  
QY 843 GGGGAGAGTAGAGCTGAAAATCCAAAGAGGTGGGGAGCCGTATGCCACCAATAAGTGAA 902  
|||||  
Db 432 AGGCGAGTAGAAGTGCATTCTGGAGAAGCTTGGATCCCACTGCTGTGATGGAACTTCAC 491  
|||||  
QY 903 CAATGCTGCAGCTGATGCTGATCAAGCAGTCTGGGATGGGAACCCGACTTCACTTCGC 962  
|||||  
Db 492 ACTTGCACACTGCCAGATCATCTGTGCAGAGTGGGTGTGGCAAGGCTGTCTGTCTCT 551  
|||||  
QY 963 TGGCTTGCCTCATTTGCACTCAGGCTCTGATGTTGTATGGCTTGTATGCTCTCTCTCT 1022  
|||||  
Db 552 GGGACATGAGGCTCTTCAAGAGTCCAGTCCCAAGCTTGGGCTGAAGAGTTCAGGTGGA 611  
|||||  
QY 1023 CGGTAATGAATCTTTCTTTGGGACTGCAGACATTTCCGGAACCCGTCATTTTGACTGTCT 1082  
|||||  
Db 612 GGGGAGGAGCTCAGCTCTGGGTCTGCCCAAGAGTGCCTGTCCAGGGGCGACGTGTCA 671  
|||||  
QY 1083 TCATCAAAACGATGTGCTGTGATGCTCTCAGATGGAGCAGATTTGGAACTGCCACTGAC 1142  
|||||  
Db 672 CCACAGTGGATCTGCTCAGGTTGTTTGTTCAGCATACTCAGA---AGTCCGGCTCATGAC 728  
|||||  
QY 1143 AGATGGAAGTAACAATTTCTTCAGGAGAGTAGAGTGAATTCATGAACAGTGTGGAC 1202  
|||||  
Db 729 AACGGCTCTCTCAGTGTGAAGGCGAGTGGAGTGAACATTTCTGGACAATGAGGAGC 788  
|||||  
QY 1203 AATATGTACCAAGTGAAGAATGAACAGCCCTTGTGCTTTGTAAAGAGCTAGGATG 1262  
|||||  
Db 789 GCTCTGCTCTCCACTGGAGTCTGGCCAATGCCAATTTATCTGCTCAGCTCGGCTG 848  
|||||  
QY 1263 TCCGTTTCAAGCTCTTTGGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1322  
|||||  
Db 849 TGGAGTTGCCATCTCCACCCCGGAGGACACACTTGGTGAAGAGGATGATCAGATCCT 908  
|||||  
QY 1323 GATAACAGCATATCTTGCACCTGGGAATGAGTCACTCTCTGGGACTGCACATATGATGG 1382  
|||||  
Db 909 AACAGCCGATTTCACTGCTCTGGGCTGAGTCTCTCTGAGGAGTCTCTCTGCTGCTGCTG 968  
|||||  
QY 1383 AAAAGCAAGGCAACATGCTTCCGAAGATCAGATGCTGGAGTAAATTTGCTGTGTAAGGC 1442  
|||||  
Db 969 CCTGGTGGTCTCTGACTGTTCCCATGGCAACACAGCCTCTCTGATCTCTCAGGAACCA 1028  
|||||  
QY 1443 AGATCTGG----- 1450  
|||||  
Db 1029 GATCCAGGTGCTTCCCAGTGAACAGCTCGCTGTCTCAACCTACAGGCTCTGCGGCTC 1088  
|||||  
QY 1451 -----ACCTAAGGCTGTGCGGGCTCATAG 1475  
|||||  
Db 1089 AGAGACAGCCGCCCTTACTCTCTCAGACAGGAGGCTCCGCTGCTGGAGCGGGGGG 1148  
|||||  
QY 1476 CCCCTGTTATGGAGATGGAGGTGAATACCAAGAGAGTGGGGAGTGTGCTGCTCATGA 1535  
|||||  
Db 1149 TCCCTGGCCGGAGAGTGGAGATCTTGGACAGGCTCTCTGGGCACTCTGTGATGA 1208  
|||||  
QY 1536 CAGATGGAGCACAGGAATGAGCTGTTGTGTAAACAATTTGGATGTGGAAGACCTAT 1595  
|||||  
Db 1209 GGGTGGGACCTGGACGATGCCCGCTGTGTGAGGAGCTGGGCTGTGAGAGAGCCCT 1268  
|||||  
QY 1596 GCATGTGTTGCTATGACCTATTAAAGACATCAGGACCTATTTCGGCTGGATGAGCT 1655  
|||||  
Db 1269 CAATGCCACGGGCTCTGCTCACTTCGGGGCAGGATCAGGGGCCCATCTCTGTTGGACA 1328  
|||||  
QY 1656 TTCTTGCATTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGTGATGGGAAAGCA 1715  
|||||  
Db 1329 GAATTCACAGAGAGGAGTCCACGCTGTGAGGTGCCCTTCCCGGGGCTGGGGCAGCA 1388  
|||||  
QY 1716 TAATTGTGTACACAGAGAGGATGTGATTTAACTGTCTCAGGTGATGCAACATGGGGCT 1775  
|||||

Db 1389 CAACATGACAGACAAAGCAGGAGCGGGGGTCACTCTGCTCAGAGT-----TCCTGGCCCT 1442  
Qy 1776 GAGGCTGTGGCGGCGAGCAACCGCTGCTGGGAAGACTGAGGTGTACTTTCAAGAGCG 1835  
Db 1443 CAGGATGGTGAATGAGGACCAAGTGTGCTGGGTGGCTGGAAGTTTCTACAAATGGGAC 1502  
Qy 1836 GTGGGACACAGTGTGTATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGAGCCA 1895  
Db 1503 CTGGGACAGTGTGCTGCGGTAAACCCATGGAAGACATCACTGTGTCCACGATCTGCAGACA 1562  
Qy 1896 GCTGACATCCCATCTTCTATCATTTGGCATGGGTCTGGGAACGCTTCTACAGGATATGG 1955  
Db 1563 GCTTGGCTGTGGGACAGTGGAAACCTCAACTCTTCTGTGCTCTAGAGAAGGTTTTAG 1622  
Qy 1956 AAAAATTTGGCTGCATGATGTTTCTGTGATGGATGAGTCAAGTCTCTGTGCTATGCG 2015  
Db 1623 GCCAAGTGGGTAGTAAGTCCAGTGTGGGAACCTGACACCTCTCTGTGCAAGTGTCC 1682  
Qy 2016 GAACAGTGGGTGGGAATTAATGACTGACAGTCAACAGTCAAGATGTTGGAGTGTCTGTT 2075  
Db 1683 TTCTGACCCCTTGAATTACAACCTATGCTCTCCAAAGGAGGAGCTATATCTGGTG--- 1739  
Qy 2076 TGATGATCGGATATGGAGCTGAGCTGTGGTGAACGACAGCTGTGCTGGAAGT 2135  
Db 1740 ---TGACAGACAGACAGATCCGCTGTGGTGGATGGAGTGTGCTCTGGAGAGT 1796  
Qy 2136 TGAGGTGAATGCCAGGTGCGGTGGGAATCTGTGTAAATGGCTGGGGAATGAACAT 2195  
Db 1797 GGAGATCCTTGACAGGCTCTGGGACCACTGTGTATGACCGTGGGACCTGGACGA 1856  
Qy 2196 TGGTGAAGTGTGACAGCACTTGAATGTGGGTCTGCAATCAGGCTTCCAGAGGCC 2255  
Db 1857 TGCCCGTGTGGTGTGAAGCAGTGGGCTGTGGAGAAGCCCTGGACGCCACTGTCTCTTC 1916  
Qy 2256 TCATTTACAGAAAGAACATACACATCTTAATGTGCAATTTCTGGCTGCAGTGGAGGGA 2315  
Db 1917 CTCTCTGGGACGGGATCAGGCCCTCTGGCTGGATGAAGTGAAGTGAAGTGAAGGAGAGA 1976  
Qy 2316 AGCCTCTCTGCGGATGTATAGATGGAGTGGAGTGGAAACAGAGCTGCGTGTCAATTAAT 2375  
Db 1977 GTCCCAAGTATGAGGTGCCCTTCTGGGATGGGCGGCAACAACTGCAATCATCAAGA 2036  
Qy 2376 GGAGCAAGTTGTATGCTCAGCCACAGGAGCCAGGCTGGTGGAGCTGATATGCC 2435  
Db 2037 AGATGAGGAGTCACTGTCTAGGATTTGTGC-----GTCTGGCTGGAGGATGGACC 2090  
Qy 2436 CTGCTCTGAGCTGTGGAAGTGAACATCAGACATCGCGCTCTGTCTGTGATTCTCA 2495  
Db 2091 CTGCTCAGGGCAGTGAAGTGAATTTGGAGAAGCTTGACCCCAAGTGTCTGTGGAAGA 2150  
Qy 2496 TTTCTCTCTTCTGCTGCCAATGTGCTGCAGAGAAATTAATTTGTGGAGATGCCATATC 2555  
Db 2151 CTTCACACTCCCACTGCCAGGTCTATCTGTGCAGAGCTGGGATGGCAAGGCTGTGTC 2210  
Qy 2556 TCTTTCTGTGGAGATCACTTTGGAAAGGAATGTCTTAATTTGGGCCGAAAGTTCCA 2615  
Db 2211 TGTCTCTGGACACATGCCATTCAGAGAGTCCGATGGCCAGGCTGTGGGCTCAAGAGTTCA 2270  
Qy 2616 GTGTGAAGGAGTGAACATCACTTGCATATGCCCATTTGTTCAACATCCGGAAGACAC 2675  
Db 2271 GTGTGATGGGGGAGCCCTGAGTCTGTGGCTGCCCAAGTGGCTGTCCAGGAGGCAC 2330  
Qy 2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTCTCCGATATACAGATGTCCGACTGT 2735  
Db 2331 ATGCTCTCCAGTGGAGCTGCTCAGGTGTCTGTCTCAGTGTACACAGAAGTCCAGCTTAT 2390  
Qy 2736 G---ANTGGCAATCCAGTGTGACGGCAAGTGGAGATCAACGCTTTGGACACTGGGG 2792  
Db 2391 GAAAAACGGCACCTCTCAATGTGAGGGCAGGTGGAGATCAAGATCTCTGGACGATGGAG 2450  
Qy 2793 CTCACTGTGTGACACCCACTGGGACCCAGAGATGCCGTGTCTATGACAGAGCTCAG 2852  
Db 2451 AGCGCTCTGTGGCTCCCACTGGAGTGTGGCAATGTCATGTCCTGTCTGTCTGTCTGTCTCG 2510

Qy 2853 CTGTGGGACTCTCTCTCAACACAGAGAGAAATATATTTGGAGAAAGAGTTCTGTGT 2912  
Db 2511 CTGTGGAGTGCCTACTCCACCCAGAGGACCACTTGTGTGGAGAGGTGATCAGAT 2570  
Qy 2913 GTGGGACACAGGTTTCAATTTGCTTAGGAATGAGTCACTTCTGGGATAACTGTCAATGAC 2972  
Db 2571 CTCACACCCCAATTTCACTGCTCAGGGCTGAGTCTTCTCTGGAGTTGTCTGTGTGAC 2630  
Qy 2973 AGTTCTTGGAGCACCTCCCTGTATCCATGGAAATACTGTCTGTGATCTGCACAGGAAG 3032  
Db 2631 TGCTTGGTGGGCTGACTGTTCCATGGCAACAGCCTCTGTGATCTGCTCAGGAAA 2690  
Qy 3033 COTGACCCAGCACCTGTTTCCATGCTCGCAATATATCTGACCCATATTTGTCTGCAAGT 3092  
Db 2691 CCACACCCAGGTGCTGCCCGGAGTGCACAGACTTCTGTCTCAACCTGCAGGCTCTGCGGC 2750  
Qy 3093 TCCAGAGGCGAGTGTCTTGTATCTGCTTAGAGGACAAACGGTCCCGCTAGTGGATGGGA 3152  
Db 2751 CTCAGAGGAGTTTCTCCCTACTGCTCAGACAGCAGGAGCTTCCCGCTGGTGGAGCGGG 2810  
Qy 3153 CAGCGCTGTGCGGAGAGTAGAGTCTATCACCAGCGCTTCTGGGCGACCATCTGTGA 3212  
Db 2811 CGTCTCTGCGGCGGAGAGTGGAGTCTTGACCAAGGCTCTCTGGGCGACCATCTGTGA 2870  
Qy 3213 TGACGGCTGGGACCTGAGCGATGCCAGTGGTGTCTCAAAAGCTGGGCTGTGGAGTGGC 3272  
Db 2871 TGATGACTGGGACCTGGAGTGGCGTGTGTGTGTCAGAGCAGCTGGGCTGTGGAGAAGC 2930  
Qy 3273 CTTCAATGCCAGGCTCTGCTCACTTTGGGAGGGGTGAGGCCCATCTGTGCTGATCA 3332  
Db 2931 CCTCAATGCCAGGGGTCTGCTCACTTGGGCGAGATCAGGGCCCATCTGTGCTGACGA 2990  
Qy 3333 COTGAACCTGCACAGGAAGGAGTCCCACTTGTGGCAGTGCCTTCCCGGGCTGGGGGCA 3392  
Db 2991 COTGAACCTGCACAGGAAGGAGTCCCACTGTTGGAGTGCCTTCCCGGGCTGGGGGCG 3050  
Qy 3393 GCACACTGCAGGACACAGGAGGAGCGAGGCTCATCTGCTCAGAAATTCACAGCCTTGAG 3452  
Db 3051 GCACACTGCAGACACAAGGAGGAGCGCGGCTCATCTGCTCAGAGTTCTCTGGCCCTCAG 3110  
Qy 3453 GCTCTACAGTGAACCTGAAACAGAGAGCTGCTGCTGGAGATTGGAAGTCTCTATAACGG 3512  
Db 3111 G-----ATGGTGAAGGAGGAGGAGTGTGCTGGTGGCTGGAGGTTTCTACACGG 3164  
Qy 3513 GACCTGGGCGAGCTGCGGAGGAGGAACATCACACAGCAGATAGCAGGCAATGTGTGACAG 3572  
Db 3165 GACCTGGGCGAGTGTGCGGCGAGCCCATGGAAGATATCACTGTGCTCGTGTATGTCAG 3224  
Qy 3573 GCAGCTGGGCTGTGGGGAATGAGTGTGACGCTCGCCCTTTATCTAAGACAGGCTC 3632  
Db 3225 ACAGCTTGGATGTGGGCGACAGTGAAGTCTCAACACTCTGTTGGTCTCAGGGAAGGTTTC 3284  
Qy 3633 TGGTTTCAGTGGGTGATGACATTCAGTCTTAAACGCAATATCTCCATATGGCAGTG 3692  
Db 3285 TAGACCCCGGTGGGTAGATTTAATTCAGTGTGGAATAATGAGATACCTCTCTCTGGCAGTG 3344  
Qy 3693 COTGCTGCCCCATGGGCGGAAGATCTCCAGCCCGCAGAGAGACCTGGATCACATG 3752  
Db 3345 TCTTCTGCCCCATGGAATACAGTTCAGTCTCTCAAGAGGAGGAGCCTACATCTCATG 3404  
Qy 3753 TGAAGATAGATAA----- 3766  
Db 3405 TGAAGGAAGAAGACCCAGAGCTGTCCAACTGCTGCGGCTGCACAGACAGAGAGAGCT 3464  
Qy 3767 ---GAGTGTGTGAGGAGACACAGTGTCTGGAGAGTGTGAGATCTGCGACGAGGCTC 3824  
Db 3465 CCGGCTCTAGGGGAGGAGACAGCGGAGTGTCTCAGGGGTGGAGGTGTGTCACACAGGCTC 3524  
Qy 3825 CTGGGCGACAGTGTGTGATGACTCTCTGGGACCTGGCGGAGGAGGAGTGTGTGTGACGA 3884  
Db 3525 CTGGGCGACCGTGTGGATGACTCTCTGGAGCCTGGCAGAGGCTGAGGTGTGTGTGACGA 3584



Db 1149 TCCCTGCGCCGGAGTGGAGATCCTTTGACCAGGGCTCTCGGGCACAATCTGTGATGA 1208  
Qy 1536 CAGATGGAGCACAAGGAATGACAGCTGTTCTGTGTAAACAATTTGGATGTGGAAAGCCAT 1595  
Db 1209 CGCTGGGACCTTGGAGATGCCCGGTGTGTGACAGGAGCTGGGCTGTGGAGAAGCCCT 1268  
Qy 1596 GCATGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGAGCT 1655  
Db 1269 CAATGCCAGGGGTGCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGGTTGGACAACCT 1328  
Qy 1656 TTCTTGCATTTGAATGAGTCAAAATCTGGGACGTGTGAACACAGTGGATGGGAAGCA 1715  
Db 1329 GAACCTGACAGAAAGAGTCCACGTGTGGAGGTGCCCTCCCGGGCTGGGGCAGCA 1388  
Qy 1716 TAATTGTGTACACAGAGAGTGTGATTTAACTCTGCTCAGGTGATGAACATGGGGCT 1775  
Db 1389 CAACTGCACACAAAGCAGCGGGGTCTCTCTCAGAGT-----TCTGGCCCT 1442  
Qy 1776 GAGGTGTGGGGGAGCAAGCCGTCTCGGAGAGACTGGAGGTGTACTTTCAAGGAGC 1835  
Db 1443 CAGGATGGTGTGAGTGAGGACAGCAGTGTGCTGGGTGGAGTGTTCACAATGGGAC 1502  
Qy 1836 GTGGGSCACAGTGTGTGATGAGCGGTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCA 1895  
Db 1503 CTGGGCGAGTGTCTGCCGTAAACCCATGGAACACATCACTGTGCCAGATCTGCAGACA 1562  
Qy 1896 GCTGGAAGTCTCCCTCTCTATCACTTGGCATGGGTCTGGGAAACGGTCTTACAGGATATGG 1955  
Db 1563 GCTTGGCTGTGGGACAGTGAACCTCAACTCTTCTGTGCTTTAGAGAAGTTTTAG 1622  
Qy 1956 AAAAATTTGGCTGATGATGTTTCTGTGATGAGATGAGTCAAGTCTCTGTGTCATGCAAG 2015  
Db 1623 GCCACAGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682  
Qy 2016 GAACAGTGGGTGGGAAATCACTGACAGTCACTGACAGTGAAGATGTTGGAGTATCTGTT 2075  
Db 1683 TTCTGACCTTGAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1739  
Qy 2076 TGATGATCGGATATGGAGTGTGAGTGTGGGTGGAAGCAGCAGTGTGCTGGAAAGT 2135  
Db 1740 ---TGACAGACAGACAGATCCGCTGTGGATGGAGTGTGCTGTCTGGGAGAT 1796  
Qy 2136 TGAGTGAATATCCAGGTTGCCGTGGGAATTTCTGTGTGATGATGATGATGATGATGATGAT 2195  
Db 1797 GGAGATTCCTTGACAGGGTCTTGGGGACCATCTGTGATGATGATGATGATGATGATGATGAT 1856  
Qy 2196 TGTGAAGTGTGTTGAGGCAACTTGAATGTGGTCTGCAATCAGGCTTCCAGAGAGCC 2255  
Db 1857 TGCCCGTGTGGTGTGACAGAGTGGCTGTGGAGAGCCCTGGACGCCACTGTCTCTTC 1916  
Qy 2256 TCATTTTCAGAAAGAACATTAACATCTTAATGTGGAATTTGGCTGCACTGGAGGGGA 2315  
Db 1917 CTCTTCGGGAGGGATCAGGGCCATCAGGGCTGATGAAGTGAATGCAATCATCAAGA 1976  
Qy 2316 AGCCTCTCTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2375  
Db 1977 GTCCCAAGTATGGAGTGGCTTCTTGGGGATGGGGCAACAACTGCAATCATCAAGA 2036  
Qy 2376 GGAAGCAAGTTTGTCTGTCTGACCCACAGGAGCCAGGCTGGTGGAGCTGATATGCC 2435  
Db 2037 AGATGAGGAGTCACTGTCTCAGGATTTGTG-----GTCTGGCTGGAGGATGGACC 2090  
Qy 2436 CTGCTCTGAGCTGTTTGAAGTGAACATCAGACATAGCAGTGGCTCTGTCTGTGATTTGA 2495  
Db 2091 CTGCTCAGGCGAGTAGAAGTGCATTTCTGGAAAGCTGGACCCAGTGTCTGTATGAAA 2150  
Qy 2496 TTCTCTCTTCATGCTGCGCAATGTGCTGTGACAGAGATTTAAATTTGGAGATGCCATATC 2555  
Db 2151 CTTACACTCCCCACTGCCCCAGGTCTATCTGTGACAGCTGGGATGGCAAGGCTGTGTC 2210  
Qy 2556 TCTTTCTGTGGGAGATCACTTTTGAAGGAGTGTCTTAACTTGGCCCGCAAAAGTTCCA 2615  
Db 2211 TGTCTGGGACATGCCATTCAGAGAGTCCGATGCCAGGCTCGGCTGAGAGTTGAG 2270

Qy 2616 GTGTGAAGGAGTGAACCTCACTTGCATATATGCCCATTTGTTCAACATCCGGAAGACAC 2675  
Db 2271 GTGTGATGGGGGAGCCTGAGCTCTGTGCTGCCAGAGTCCCTGTCCAGGAGGCAC 2330  
Qy 2676 TTGTATCCACACAGCAGAGAAGTTGGAGTTGTCTGTGTTCCCGATATACAGATGTCGACTTGT 2735  
Db 2331 ATGTCTCCACAGTGGAGTGTCTGTTTCTGTCAGTGTACACAGAATCCAGCTTAT 2390  
Qy 2736 G---AATGSCAAATCCAGTGTGACGGGCAAGTGGAGATCAACGTCTTGGACACTGGGG 2792  
Db 2391 GAAACAGGCACCTCTCAATGTGAGGGGAGGTGGAGATGAAGATCTCTGGACGATGGAG 2450  
Qy 2793 CTCACTGTGTACACCACTTGGACCCAGAGATGCCGTGTCTATGACAGACAGTCTAG 2852  
Db 2451 AGCGCTGTGTGCTCCACTGCGAGTCTGGCAATGCAATGTTGTCTGTCGACGCTCG 2510  
Qy 2853 CTGTGGGACTGCTCTCAACACAGGAGGAAATATATTTGGAGAAAGTGTTCGTCT 2912  
Db 2511 CTGTGAGTGGCCATCTCCACCCAGAGGACCACTTGGTGGAGGAGGTGATCAGAT 2570  
Qy 2913 GTGGGACACAGGTTTCAATTTAGGGAATGAGTCACTTCTGGATAACTGTCAATGAC 2972  
Db 2571 CTCACAGCCCAATTTCACTGCTCAGGGCTGAGTCTTCTCTGAGTGTCTCTGTGAC 2630  
Qy 2973 AGTTCTTGGAGCACCTTCCCTGTATCCATGGAATACTGTCTGTGATCTGCACAGGAAG 3032  
Db 2631 TGCTTGGTGGGCTGACTGTTCCATGGCAACACAGCTCTGTGATGCTGTGTCAGGAAA 2690  
Qy 3033 CCTGACCCAGCACGTTTCCATGCTCCCAATGTATCTGACCCATATTTGTCTCAGT 3092  
Db 2691 CCACACCCAGTGTGCTGCCAGTGCACACTTCTCTCAACCTGACGGCTCTCGGG 2750  
Qy 3093 TCCAGAGGCACTGCTTTGATCTGTAGAGCAAAACGGCTCCGCTAGTGGATGGGA 3152  
Db 2751 CTCAGAGAGATTTCTCCCTACTGCTCACACAGGAGCTCCGCTGTGGACGGGG 2810  
Qy 3153 CAGCCGCTGTGCGGGAGAGTATAGATCTATCAGCAGCGCTTCTGGGGCACCATCTGTGA 3212  
Db 2811 CGTCCCTCGGGGAGAGTGGAGATCTTGAACAGGCTCTCTGGGGCACCATCTGTGA 2870  
Qy 3213 TGACGCTGGGACCTCAGCGATGCCACGCTGTGTCTCAAAAGCTGGCTGTGGAGTGC 3272  
Db 2871 TGATGACTGGGACCTGGAGATGCCGCTGTGTGTGTCAGGAGCTGGCTGTGGAGAGC 2930  
Qy 3273 CTTCATGTCAGGCTCTCTGCTCACTTTTGGGAGGGGTGAGGGCCCATCTGGCTGTGATGA 3332  
Db 2931 CCTCAATGCCAGGGTCTGCTCACTTCGSGGAGGATCAGGGCCCATCTGGCTGGACGA 2990  
Qy 3333 CCTGAATGCACAGGAGCGGATCCCACTTTGGCAGTGCCTTCCCGGGCTGGGGGCA 3392  
Db 2991 CCTGAATGCACAGGAGGATGCCACGCTGTGGAGTGCCTTCCCGGGCTGGGGGG 3050  
Qy 3393 GCAGACTGCAGGCACAAGGAGGAGCAGGGTCACTGCTCAGAAATTCACAGCCCTTGAG 3452  
Db 3051 GCAGACTGCACACACAGAGAGGAGCCGCGGTCTACTGCTCAGAGTTCTGTGGCCCTCAG 3110  
Qy 3453 GCTCTACAGTGAACATGAACAGAGAGCTGTGCTGGAGATTGGAACTTCTTATTAACGG 3512  
Db 3111 G-----ATGGTGAAGGAGGAGGAGTGTGCTGGTGGCTGGAGGTTTCTTACAACGG 3164  
Qy 3513 GACCTGGGCGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3572  
Db 3165 GACCTGGGCGAGTGTGCTGCCAGGCCCATGGAAGATATCACTGTGCTGCTGATCTGCGAG 3224  
Qy 3573 GCAGCTGGGCTGTGGGAGAAATGGAGTGTGCTGAGCCCTGCGCCCTTTATCTAAGACAGGCTC 3632  
Db 3225 ACAGCTTGTGATGGGAGCAGTGAAGTCTCAACACTCTGTGGTCTCAGGGAAGGTTTC 3284  
Qy 3633 TGTGTTTCTGATGGGAGATCAGATTCAGTGTCTTAAACGACATATCCATATGCGAGTG 3692  
Db 3285 TAGACCCCGGTGGGTAGATTTAATTCAGTGTGCGGAAATGGATACCTCTCTCTGCGAGTG 3344



Db 391 TGGATGACATCGCGTCAAGGAAATGAGTCAITTTCTATGGGACTGTACGCCAACCCCT 450

QY 4016 GGGACAGAGTACTGTGGACA 4037

Db 451 GGGACAGAGTACTGTGGACA 472

## RESULT 9

US-10-040-739-138

; Sequence 138, Application US/10040739

; Patent No. US20020173635A1

; GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth

; McCoy, John

; LaValle, Edward

; Racie, Lisa

; Merberg, David

; Treacy, Maurice

; Spaulding, Vikki

; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS

; NUMBER OF SEQUENCES: 1519

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genetics Institute, Inc.

; STREET: 87 CambridgePark Drive

; CITY: Cambridge

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02140

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/040,739

; FILING DATE: 07-Jan-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/036,520

; FILING DATE: 03-JUN-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Brown, Scott A.

; REGISTRATION NUMBER: 32,724

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 498-8224

; TELEFAX: (617) 876-5851

; INFORMATION FOR SEQ ID NO: 138:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 599 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 138:

US-10-040-739-138

Query Match 4.6%; Score 213; DB 9; Length 599;

Best Local Similarity 68.0%; Pred. No. 4.6e-48;

Matches 297; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 1770 GGGCGTGGCGTGGGGGGGCGAGCAACCGCTGCTCGGGAACACTGGAGGTCTACTTTCA 1829

Db 161 GGAGCTGGAGCTAGTGGTGTGAAACAAAGTGTAGCGGGGAGAGTGGAAAGTGAAGTCCA 220

QY 1830 AGGACGGTGGGGCACAGTGTGTGATCACGGCTGGAACTGAAGCTGCAGCTGTGGTGTG 1889

Db 221 GGAGGAGTGGGAACCGTGTGTAATAATGCTGGAGCATGGAAGCGGTCTCTGTGATTG 280

QY 1890 TAGCCAGCTGGAGTCCCACTCTCTATCATATTGCGATGGGTCTGGGAAACGCTTCTACAG 1949

Db 281 TAACCAAGCTGGGATGTCCCAACTGCTATCAAGCCCGCTGGATGGGCTTAATTCAGTGCAGG 340

QY 1950 ATATGAAAAAATTGGCTCGATGATGTTTCTCTGTGATGGAGATGAGTCAGATCTCTGGTC 2009

Db 341 TTCTGGACGCAATTGGATGATCATGTTCTTGTCTGGGAATGATCAGCTCTTTGGGA 400

QY 2010 ATGCAAGAACAGTGGGTGGGAAATATATGACTGCAGTGCAGTCAAGATGTTGGAGTGTAT 2069

Db 401 TTCCAAACATGATGGATGGGGAAGCATTTATGCAATCAACAATGAAGATCTGGCGTGAC 460

QY 2070 CTGTTCTGATCGATCGGATATGAGCTGAGGCTTGTGGGTGGGAAGCAGCAGGTGTGCTGG 2129

Db 461 ATGTTCTGATGGATCAGATCTGGAGCTTAAGACTTAGAGGTGGAGCGCGCTGTGCTGG 520

QY 2130 ARAAGTTGAGTGAATGTCAGGCTGCCGTGGGAATTTCTGTGCTAATGCTGGGGAAT 2189

Db 521 GACAGTTGAGTGGAGATTCAGAGACTGTTAGGGAAGAGTGTGTGCACAGAGCTGGGACT 580

QY 2190 GAACATTGCTGAAGTTG 2206

Db 581 GAAAGAAGCTGAAGCTG 597

## RESULT 10

US-09-796-692-269

; Sequence 269, Application US/09796692

; Publication No. US20020198362A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

; FILE REFERENCE: 2077,001200

; CURRENT APPLICATION NUMBER: US/09/796,692

; PRIOR FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: 60/218,950

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/222,903

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: 60/223,416

; PRIOR FILING DATE: 2000-08-04

; PRIOR APPLICATION NUMBER: 60/223,378

; PRIOR FILING DATE: 2000-08-07

; NUMBER OF SEQ ID NOS: 9597

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 269

; LENGTH: 461

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-796-692-269

Query Match 4.3%; Score 199.2; DB 9; Length 461;

Best Local Similarity 66.7%; Pred. No. 2.4e-44;

Matches 299; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

QY 183 TGGAGACGGTCCCTCTCTGGGACAGTGGAGGTGAAATTCACAGGACAGTGGGGGACTGT 242

Db 5 TGGAGGGAATATGTGTTCTGGAAGATAGAGATCAAAATTCAGAGCGGTGGGGAACACT 64

243	QY	GTGTGATGATGGGTGGAACTACTGCTCACTGTCGTGTGCAACACAGCTTTGGATGCC	302
65	Db	GTGTGATGAATCTCAACATAGATCATGCATCTGCTCATTTGTAGACAACTTGAATGTGG	124
303	QY	ATTTTCTTTGCGCATCTTTTCGTTTTTGGACAAGCCGTGACTAGACATGGAAGAAATTTGGCT	362
125	Db	AAGTGTCTGCAGTTCTCTGGTTTCATCTAATTTGGAGAAGCTCTGGACCACATCTGGTT	184
363	QY	TGATGATGTTTCTCTGTTA-TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCCGGGAAT	421
185	Db	TGATGATCTTATATGCAACCGGAATGAGTCAGCTCTGGAACTTGCAACATCAAGGAT	244
422	QY	GGGGAAGCCATAACTGTTATCATGGAGAAGATGTTGGTGTGAACCTGTTATGCTGAAGCCA	481
245	Db	GGGGAAGCATAACTGTGATCATGCTGAGGATGCTGGAGTGATTTGCTCAAAGGGAGCAG	304
482	QY	ATCTGGGTTTGAGGCTAGTGGATGGAAACAACCTCCTGTTACGGGAGAGTGGAGGTGAAT	541
305	Db	ATCTGAGCCCTGAGACTGGTAGATGGAGTCACCTGAATGTTTCAGGAAGATTTAGAAGTGAGAT	364
542	QY	TCCAAGAAAGGTGGGGAGCTATATGTGATCATGGGTGGAAACTTTGAATACTGCTGCCGTGG	601
365	Db	TCCAAGAGNATGGGGACAATATGTGATGACGGCTGGGACAGTTCACGATGCTGCTGTGG	424
602	QY	TGTGCAGGCAACTPAGGATGCCATCTTC	629
425	Db	CATGCAAGCAACTGGGATGTCCAACATGC	452

RESULT 11  
US-09-796-692-7354  
; Sequence 7354, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7354  
; LENGTH: 461  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-7354

Query Match	4.3%;	Score 199.2;	DB 9;	Length 461;
Best Local Similarity	66.7%;	Pred. No. 2.4e-44;		
Matches 299;	Conservative 0;	Mismatches 148;	Indels 1;	Gaps
Qy 183	TGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATCCAGGACAGTGGGGAGCTGT	242		
Db				
Db 5	TGGAGGAATATGTGTTCTGGAGAATAGAGATCAAAATCCAGGACGGTGGGGAACAGT	64		
Qy 243	GTGTGATGATGGGTGGGAACACTACTTGCCCTCAACTGTGCTGTGCAACACAGCTTCGGATGTCC	302		
Db				
Db 65	GTGTGATGATAAATTCACATAGATCATGCAATCTGTCAATTTGTAGACAACCTCAATATGTG	124		
Qy 303	ATTTTCTTTCCGCATGTTTCGTTTTTGGACAGCCGTGACTAGACATGGAAGAAATTTTGGCT	362		
Db				
Db 125	AGTGTGTGTCAGTTTCTCGTTCATCTAATTTTGGAGAAGGCTCTGGACCAATCTGGTT	184		
Qy 363	TGATGATGTTTCCCTGGTTA-TGGAAATGAGTCAGCTCTCTGGGAATGTCAACACCCGGAAT	421		
Db				
Db 185	TGATGATCTTATATGCAACCGGAAATGAGTCAGCTCTCTGGAACCTGCAACATCAAGGAT	244		
Qy 422	GGGAAGCCATAACTGTTATCATGGAGAAGATGTTGTGTGAACCTGTTATGGTGAAGCCA	481		
Db				
Db 245	GGGAAGACATAACTGTGATCATGCTGAGGATGCTGAGTGATTTGCTCAAAGGGAGCAG	304		
Qy 482	ATCTGGGTTTGAGGCTAGTGGATGGAGAAACAACTCCTGTTCCAGGAGAGTGGAGGTGAAT	541		
Db				
Db 305	ATCTGAGCCTGAGACTGGTAGATGGAGTCACTCAATGTTCCAGGAAGATTAGAAGTGAGAT	364		
Qy 542	TCCAAGAAAGTGGGGACATATATGATCATGGTGGGAACCTTCATATCTGCTGCCGTGG	601		
Db				
Db 365	TCCAAGGAATGGGGACAAATATGTGATGACGCTGGGACAGTTACGATCTGCTGTGG	424		
Qy 602	TGTCAGGCAACTAGAGTGTCCATCTTC	629		
Db				
Db 425	CATCAAGCAACTGGGATGTCCTCACTGC	452		

RESULT 12 .  
 US-10-040-862-269  
 : Sequence.269, Application US/10040862  
 : Publication No. US20030078396A1  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Gaiger, Alexander  
 : APPLICANT: Algate, Paul A.  
 : APPLICANT: Mannion, Jane  
 : APPLICANT: Retter, Marc  
 : APPLICANT: Corixa Corporation  
 :  
 : TITLE OF INVENTION: Compositions and Methods for  
 : TITLE OF INVENTION: Hematological Malignancies  
 : FILE REFERENCE: 014058-013520US  
 : CURRENT APPLICATION NUMBER: US/10/040,862  
 : CURRENT FILING DATE: 2001-11-06  
 : PRIOR APPLICATION NUMBER: US 60/186,126  
 : PRIOR FILING DATE: 2000-03-01  
 : PRIOR APPLICATION NUMBER: US 60/190,479  
 : PRIOR FILING DATE: 2000-03-17  
 : PRIOR APPLICATION NUMBER: US 60/200,545  
 : PRIOR FILING DATE: 2000-04-27  
 : PRIOR APPLICATION NUMBER: US 60/200,303  
 : PRIOR FILING DATE: 2000-04-28  
 : PRIOR APPLICATION NUMBER: US 60/200,779  
 : PRIOR FILING DATE: 2000-04-28  
 : PRIOR APPLICATION NUMBER: US 60/200,999  
 : PRIOR FILING DATE: 2000-05-01  
 : PRIOR APPLICATION NUMBER: US 60/202,084  
 : PRIOR FILING DATE: 2000-05-04  
 : PRIOR APPLICATION NUMBER: US 60/206,201  
 : PRIOR FILING DATE: 2000-05-22  
 : PRIOR APPLICATION NUMBER: US 60/218,950  
 : PRIOR FILING DATE: 2000-07-14  
 : PRIOR APPLICATION NUMBER: US 60/222,903  
 : PRIOR FILING DATE: 2000-08-03

```
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 269
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-269

Query Match
Best Local Similarity 4.3%; Score 199.2; DB 9; Length 461;
Matches 299; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

QY 183 TGGAGACGGTCCCTGCTGGGACAGTGGAGGTGAAATTCAGGACAGTGGGGGACTGT 242
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 5 TGGAGGGAATATGTTCTGGGAAGATAGAGATCAAAATTCAGGACGGTGGGGAACAGT 64
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 243 GGTGATGATGGTGGGACACTACTGCCTCAACTGCTGTCGTCGAACACAGCTTGGATGTC 302
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 65 GTGTGATGATAACTCAACATAGATCATGCATCTGTCAATTTGTAGACAACTTGAATGTGG 124
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 303 ATTTCTTTTCGCCATGTTTCGTTTGGACAAGCCGTGACTAGACATGGAATAATTTGGCT 362
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 125 AAGTCTGTGACGTTCTCTGTTTCATTAATTTGGAGAAGCTCTGGACCAATCTGGTT 184
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 363 TGATGATGTTTCCCTGTTA-TGGAATAGTCACTCTCTGGGAATGTCAACACCGGGAAT 421
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 185 TGATGATCTTATATGCAACCGGAAATAGTCACTCTCTGGAACGCAAAACATCAAGGAT 244
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 422 GGGGAAGCCATAACTGTTATCATGAGAGAGATGTTGGTGTGAACGTTATGTGAAGCCA 481
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 245 GGGGAAGCAACTGATGATGCTGGAGATGCTGGAGTATTTGCTCAAGGGAGCAG 304
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 482 ATCTGGGTTTGAGGCTAGTGGATGGAACAACCTCCTGTTTCAGGAGAGTGGAGGTGAAAT 541
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 305 ATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTTCAGGAAGATTAGAAGTGAGAT 364
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 542 TCCAAGAAGGTGGGGACTATATGTGATGGGTGGGAACATTGAATCTGTCGCGTGG 601
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 365 TCCAAGGAGAAATGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGG 424
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 602 TGTGAGGCAACTAGGATGTCATCTTC 629
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 425 CATGCAAGCAACTGGGATGTCCAACTGC 452
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 13
US-10-040-862-7354
; Sequence 7354, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
```

```
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7354
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-7354

Query Match
Best Local Similarity 4.3%; Score 199.2; DB 9; Length 461;
Matches 299; Conservative 0; Mismatches 148; Indels 1; Gaps 1;

QY 183 TGGAGACGGTCCCTGCTCTGGGACAGTGGAGGTGAAATTCAGGACAGTGGGGGACTGT 242
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 5 TGGAGGGAATATGTTCTGGGAAGATAGAGATCAAAATTCAGGACGGTGGGGAACAGT 64
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 243 GTGTGATGATGGTGGGACACTACTGCCTCAACTGCTGTCGTCGAACACAGCTTGGATGTC 302
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 65 GTGTGATGATAACTCAACATAGATCATGCATCTGTCAATTTGTAGACAACTTGAATGTGG 124
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 303 ATTTCTTTTCGCCATGTTTCGTTTGGACAAGCCGTGACTAGACATGGAATAATTTGGCT 362
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 125 AAGTCTGTGACGTTCTCTGTTTCATTAATTTGGAGAAGCTCTGGACCAATCTGGTT 184
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 363 TGATGATGTTTCCCTGTTA-TGGAATAGTCACTCTCTGGGAATGTCAACACCGGGAAT 421
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 185 TGATGATCTTATATGCAACCGGAAATAGTCACTCTCTGGAACGCAAAACATCAAGGAT 244
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 422 GGGGAAGCCATAACTGTTATCATGAGAGAGATGTTGGTGTGAACGTTATGTGAAGCCA 481
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 245 GGGGAAGCAACTGATGATGCTGGAGATGCTGGAGTATTTGCTCAAGGGAGCAG 304
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 482 ATCTGGGTTTGAGGCTAGTGGATGGAACAACCTCCTGTTTCAGGAGAGTGGAGGTGAAAT 541
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 305 ATCTGAGCCTGAGACTGGTAGATGGAGTCACTGAATGTTTCAGGAAGATTAGAAGTGAGAT 364
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 542 TCCAAGAAGGTGGGGACTATATGTGATGGGTGGGAACATTGAATCTGTCGCGTGG 601
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 365 TCCAAGGAGAAATGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGCTGTGG 424
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 602 TGTGAGGCAACTAGGATGTCATCTTC 629
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 425 CATGCAAGCAACTGGGATGTCCAACTGC 452
Db      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 14
US-09-147-947-5
; Sequence 5, Application US/09147947A
; Patent No. US20020160490A1
; GENERAL INFORMATION:
; APPLICANT: TSURUOKA, No. US20020160490A1uo
; APPLICANT: YAMASHIRO, KYOKO
; APPLICANT: YAMAGUCHI, No. US20020160490A1omi
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease
```

FILE REFERENCE: 001560-349  
CURRENT APPLICATION NUMBER: US/09/147,947A  
CURRENT FILING DATE: 1997-03-24  
EARLIER APPLICATION NUMBER: PCT/JP98/03324  
EARLIER FILING DATE: 1998-07-24  
EARLIER APPLICATION NUMBER: JP 9/213969  
EARLIER FILING DATE: 1997-07-24  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 2562  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
OTHER INFORMATION:  
US-09-147-947-5

Query Match 4.3%; Score 197.6; DB 9; Length 2562;  
Best Local Similarity 52.2%; Pred. No. 2.1e-43;  
Matches 505; Conservative 0; Mismatches 439; Indels 24; Gaps 2;  
Qy 1133 TGCAGCTAGCAGATGGAAGTAACAATTTTCAGGAGAGTAGAGGTGAGAATTCATGAAC 1192  
Db 580 TTCGCCCTTGTGGAGGACGAGCTGTCATGAGGCCGGTGGAGCTCTACCATGCTGCC 739  
Qy 1193 AGTGGTGACAAATATGTGACAGAACTTGGAGAAATGAACAAGCCCTTGTGTTTGAAGC 1252  
Db 740 AGTGGGAACCGTTTGTGATGACCAATGGATGATGCCGATGACAGAGTGTGCGAGC 799  
Qy 1253 AGCTAGAGTGCCTTTCAGGCTCTTTGGCAGTCTGCTGTGTAACCTAGTAAGTGAAGCTA 1312  
Db 800 AGCTGGGCTCAGTGGCATGCCAAGCATGGCATACGAGTATATTTGGGAAGGCTG 859  
Qy 1313 GAGACATTTGGATAACAGCATATCTTGCAGTGGAAATGAGTCACTCTCTGGCACTGCA 1372  
Db 860 GCCCAGTTATGTTGGATGAGTACGCTGCGACTGGGAATGAGCTTCAATTGAGCAGTGC 919  
Qy 1373 CATATGATGAAAAAGCAAGCAATCTTCCGAAGATCAGATGCTGGAGTAATTTGTT 1432  
Db 920 CAAGAGCTCTCGGAGAGACATAACTGTGCCATAAAGAGATGCTGGAGTCTCTGTA 979  
Qy 1433 CTGATAGGCGAGATCTGGACCTAAGGCTTGTGGGGCTCATAGCCCTGTTATGGGAGAT 1492  
Db 980 CCCCTCTAACAGATGGGTCTATCAGATCTGCAAGTGGGAAGGAGGAGCATGAGGCTGCT 1039  
Qy 1493 TGGAGGTGAATACCAAGGAGTGGGAGTGTGTCATGACAGATGAGACAGCAAGGA 1552  
Db 1040 TGGAGGTATATTACAGAGCCAGTGGGAAGTGTCTGTGATGATGGCTGGACTGAGCTGA 1099  
Qy 1553 ATGCAGCTGTTGTGTAAACAATTTGGGATGTGAAAGCCCTATGCATGTGTTTGGTATGA 1612  
Db 1100 ATACATAGCTGTTGTGCGAGATGGGATTTAAATATGTTAAACAAGCATCTG--CCA 1156  
Qy 1613 CCTATTTAAAGAGCATCAGGACCTATTTGGCTGGAGTACGTTTCTTGCAATGGAAATG 1672  
Db 1157 ACCATTTTGAAGAAGCAGAGGCGCCATATGTTGGATGAGCTGAGCTGAGCTGAGCAAGG 1216  
Qy 1673 AGTCAATATCTGGAGCTGTGAACACAGTGGATGGGAAGCAATATGTTGACACAGAG 1732  
Db 1217 AAACAGATTTCTCAGTGTGTTCCAGGCGACAGTGGGAAGGAGTCACTGAGCCACCGCG 1276  
Qy 1733 AGGATGTGATTGTAACCTCTCAGCTGATGCAACATGGGGCTG----- 1776  
Db 1277 AAGATGTTAGCATTTGCCTGTCTACCTCGCGGAGGAGACAGAGCTCTCTCGGTTTTC 1336  
Qy 1777 -----AGCTGGTGGCGGACAGCAACCGCTCTCGCGAAGACTGGAGGTGTACTTCAAG 1831  
Db 1337 CTGTCAGATCATGATGGAGAAAATAAGAAAGAACGACGAGTGGAGTGTATCAATG 1396  
Qy 1832 GACGGTGGGCGACAGTGTGATGATGACGGCTGGAAACAGTAAAGCTGAGCTGTGTGTGTA 1891  
Db 1397 GCCAGTGGGAACAATCTGTGATGATGGATGGATGTAAGGATGCAGCTGTGATCTGTC 1456

## RESULT 15

US-09-147-947-3  
Sequence 3, Application US/09147947A  
Patent No. US20020160490A1

GENERAL INFORMATION:  
APPLICANT: TSURUOKA, No. US20020160490A1uo  
APPLICANT: YAMASHIRO, Kyoto  
TITLE OF INVENTION: No. US20020160490A1omi  
FILE REFERENCE: 001560-349  
CURRENT APPLICATION NUMBER: US/09/147,947A  
CURRENT FILING DATE: 1997-03-24  
EARLIER APPLICATION NUMBER: PCT/JP98/03324  
EARLIER FILING DATE: 1998-07-24  
EARLIER APPLICATION NUMBER: JP 9/213969  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn ver. 2.0  
SEQ ID NO 3  
LENGTH: 2614  
TYPE: DNA  
ORGANISM: Mouse  
FEATURE:  
OTHER INFORMATION:  
US-09-147-947-3

Query Match 4.2%; Score 193.2; DB 9; Length 2614;  
Best Local Similarity 52.0%; Pred. No. 3.5e-42;  
Matches 500; Conservative 0; Mismatches 438; Indels 24; Gaps 2;

Qy 1133 TGCAGCTAGCAGATGGAAGTAACAATTTTTCAGGAGAGTAGAGGTGAGAATTCATGAAC 1192  
Db 741 TTCGCCCTTGTGGTGGGAACAGTGGGCATGAAGGTGCGAGTGGAGCTGTACCACGCTGCC 800  
Qy 1193 AGTGGTGACAAATATGTACCCAGAACTGGAAGAAATGAACAAGCCCTTGTGCTTTGTAGC 1252  
Db 801 AGTGGGGAGCCATCTGTGACGACCAATGGGCAATGACAGACGACAGCTCATCTGTAGC 860  
Qy 1253 AGCTAGGATGTCCTTTCAGGCTCTTTGGCAGTCTGCTGCTTAAACCTAGTAAAGAGCTA 1312  
Db 861 AGCTGGGGCTCAGTGGCATTTGCCAAAGCATGCGATCAGGCACATTTTGGGGAAGATCTG 920  
Qy 1313 GAGACATTTGGATAAAGCAGCATATCTTGCATGGAATGATGAGTCTCTTGGGAGTGA 1372  
Db 921 GCCCAATATTGTTGATGAAGTACGCTGACCCGGAACAGCTGTCAATTGAGCAATGTC 980  
Qy 1373 CATATGATGGAAGCAAGCAACATGCTTCCGAAGATCAGATGCTGGAGTAATTTGTT 1432  
Db 981 CAAGAGCTTCTTGGGGCGCAACATAACTCTGGCCATAAAGAAAGATGCTGGAGTGTCTTGTG 1040  
Qy 1433 CTGATAAGGAGATCTCGACCTAAGGCTTGTGGGGCTCATAGCCCTGTTTATGGGAGAT 1492  
Db 1041 TTCCTCTAACAGATGGTGTCTCATCAGACTGGCAGGAGGAAAAGTACCATGAGGTGCC 1100  
Qy 1493 TGGAGGTGAATACCAAGGAGAGTGGGGAGTGTGTGTCATGACAGATGGAGCAACAAGGA 1552

Db 1101 TGGAGGTCTACTACAAAGGGGAGTGGGGGACAGTCTGTGTGATGATGGCTGGAGTGAATGA 1160  
QY 1553 ATGCAGCTGTTGTGTAAACAATTGGGATGTGGAAGCCTATGCATGTGTTTGGTATGA 1612  
Db 1161 ACACATACGTGCTGTGACGTGCTGGGATTTAAATACGGCAACA---GTCCTCTGTGA 1217  
QY 1613 CTTATTTTAAAGAAGCATCAGGACCTATTGGCTGGATGACGTTTCTTGCATTGGAATG 1672  
Db 1218 ACCATTTTGGATGGCAGCAACAGGCCCATATGGCTGGATGACGTCAGCTCAGGAAAAG 1277  
QY 1673 AGTCAATATCTGGGACTGTGAACACAGTGGATGGGAAAGCATAATTGTGTACACAGAG 1732  
Db 1278 AAGTCAGTTTCATTGAGTGTCCAGGAGACAGTGGGGAAGGCATGACTGCAGCCATAGAG 1337  
QY 1733 AGGATGTGATTGTAACTGCTCAGGTGA-----TGCAACATGGG 1771  
Db 1338 AAGATGTGGGCTCACCTGCTATCTGACAGCGATGGACATAGGCTTTCTCCAGGTTTC 1397  
QY 1772 GCCTGAGGCTGTGGCGGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTCAAG 1831  
Db 1398 CCATCAGACTAGTGGATGGAGAGAATAAGAAGGAAGGAGAGTGGAGGTTTTTGTCAATG 1457  
QY 1832 GACGTTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGTGTGA 1891  
Db 1458 GCCATGGGGAACAATCTCCGATGACGGATGGACCGATGAAGCATGCAGCTGTGATCGCC 1517  
QY 1892 GCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGCTGGGAAACGCTTCTACAGGAT 1951  
Db 1518 GGCAGCTTGGCTATAAGGCTCCTGCCAGAGCAAGGACTATGGCTTATTTTGGGGAAGGAA 1577  
QY 1952 ATGGAAAAATTTGGCTCGATGATGTTTCTGTGTGAGGATGAGTCAGATCTCTGGTCAT 2011  
Db 1578 AAGGCCCATCCACATGGGATAATGTGAAGTGCACAGGAAATGAGAAGGCCCTGGGTGACT 1637  
QY 2012 GCAGGAACAGTGGTGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTCATCT 2071  
Db 1638 GTGTCAACAAGACATTGGAGGCACAACTGCCGCCACAGTGGAGATGCAGGATCACT 1697  
QY 2072 GT 2073  
Db 1698 GT 1699

Search completed: May 12, 2003, 06:36:37  
Job time : 658 secs



Db 1546 CTTGCTGTCATGCACAGTAGTAGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGG 1605  
QY 1794 CAACCGCTCTCGGAGAGACTGGAGTGTACTTTCAAGACAGCGTGGGCACAGTGTGA 1853  
Db 1606 TGACAGGTGTACAGGCGGAGTGGAGTCTATACCAAGGCTCTCTGGGACCGGTGTGGGA 1665  
QY 1854 TGACGGCTGGAACAGTAAAGCTGACAGCTGTGTGTGTACCCAGCTGGACTGCCCATCTTC 1913  
Db 1666 TGACAGCTGGGACACCAATGATGCCAATGTCTCTGTCAGGCAACCGGCTGTGGCTGGGC 1725  
QY 1914 TATCATTTGCGATGGGTCTGGGAAACGCTTCTACAGGATATGAAATAATTTGGCTCGATGA 1973  
Db 1726 CATGTGAGCCAGGAAATGCGCGGTTTGGTCAGGCTCAGGACCAATGTCTGGATGA 1785  
QY 1974 TGTTCCTGTGATGGAGATGATGATCTCTGGTCATGCAGGACAGTGGGTGGGAAA 2033  
Db 1786 TGTGCGCTCTCTCAGGACACAGTCTTACCCGTGGAGCTGCCCCACAAATGGCTGGCTC 1845  
QY 2034 TAATGACTGAGTCAAGTGAAGATGTTGGAGTGTCTCTCTGATGC----- 2081  
Db 1846 CCACACTGTGCGCCATGTGAAGACGCTGGTGTCTCTCTGCTCAGCTTCCAGTCCCGGCC 1905  
QY 2082 -----ATCGGA 2087  
Db 1906 AACACCTAGTCCAGACACTTGGCCAACTCACATGCATCAACAGCAGGATCTGAATCCAG 1965  
QY 2088 TATGGAGCTGAGGCTTGTGGGTGGAAGCAGAGGTGTGCTGGAAGATTTGAGGTGAATGT 2147  
Db 1966 TTTGGCCCTGAGGCTGTGTAATGAGAGTGCACAGGTGTGAGGCCGAGTGGAGTCCCTATA 2025  
QY 2148 CCAGGCTGCGGTGGGAATTTCTGTGCTAATGGCTGGGGAATGAACATTTGCTGAAGTTGT 2207  
Db 2026 CCGAGGCTCTGGGCGACCGTGTGATGACTACTTGGGACACCAATGATGCCAATGTGTGT 2085  
QY 2208 TTCAGGCAACTGATGTGGTCTGCATCAGGCTCTCCAGAGCCTCATTTACACA 2267  
Db 2086 TTCAGGAGCTGGGCTGTGGCTGGGCCATGTGACGCCAGGAAATGCCCCGTTTGGCCA 2145  
QY 2268 AAGAACTATACATCTTAATGTCGAATTTGGCTGCATCTGAGGGAAGGCTCTCTCTG 2327  
Db 2146 GGGTTGAGGACCAATGTTCTCTGATGATGTGCGCTCTCAGGACATGAGTCTATCTGTG 2205  
QY 2328 GGATTTATACGATGGGAGTGGAAACAGAGTGGGTGTCTAATTAATATGAGCAAGTTT 2387  
Db 2206 GAGTGTGCCCCACAAATGGGTGGCTCTCCACAACTGTGSCCATCATGAAGAGCGTGTGT 2265  
QY 2388 GATCTGCT-----CAGCCACAGCAGGCC----- 2412  
Db 2266 CATCTGCTAGCTTCCAGTCCAGCCGACACCCAGCCAGACACTTGGCCAACTCACA 2325  
QY 2413 -----AGGCTGTTGGAGCTGATATGCC 2435  
Db 2326 TGCATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAG 2385  
QY 2436 CTGCTCTGGAGCTGTGAAGTGAACATGCAGACATGGCGCTCTGTCTGTATTTCTGA 2495  
Db 2386 GTGTGAGGCGGAGTGGAGTCTATACCGAGGCTCTCTGGGACCGGTGTGTGATGACATA 2445  
QY 2496 TTTCTCTCTCATGCTGCCAATGTGTGTCAGAGAAATTAATTTGTGGAGATGCCATATC 2555  
Db 2446 CTGGACACCAATGATGCCAATGTGTTGACAGCAGCTGGGCTGTGGTGGGCCAGCTC 2505  
QY 2556 TCTTTCTGTGGGAGATCATCTTTGAAAGGGAATGTGTTAACTTGGGCGGAAAGTTCCA 2615  
Db 2506 AGCCCCAGGAAATGCCCGTTTGGCCAGGTTTTCAGGACCAATTTGCTGATGATGTGG 2565  
QY 2616 GTGTGAGGAGTCAAACTCACTTGCATTTATGCCCATTTGTTCAACATCCCGGAAGACAC 2675  
Db 2566 CTGCTCAGGACATGAGTCTCTATCTGTGGAGCTGCCCCACCAATGGCTGTCTCCACAA 2625  
QY 2676 TTGTATCCACAGCAGAGAAGTGGAGTGTGTCTGTCTCCAGT----- 2716  
Db 2626 CTGTGGCCATCAAGACGCTGGTGTCTCTCTGCTCAGCTTCCAGTCCCGGACGACACC 2685

QY 2717 -----ATACAGA 2723  
Db 2686 CAGCCACACACTTGGCCAACTCTCGTGCATCAACAGCAGGATCTGAATCCACTTTGGC 2745  
QY 2724 TGTCCGACTGTGAATGCAAAATCCAGTGTGACGGGGAAGTGGAGATCAACGTCCTTGG 2783  
Db 2746 COTGAGACTGGTGAATGGAGGTGACAGGTGCGAGGCCGAGTGGAGSTCTTATACAAAG 2805  
QY 2784 ACATCTGGGGCTCAGTGTGTGACACCCACTGGGACCCACAAAGATGCCCTGTCTTATGAG 2843  
Db 2806 CTCCTGGGGCACCGTGTGTGATGACTTGGGACACCAATGATGCCAAGTGTGCTGACAG 2865  
QY 2844 ACAGCTCAGCTTGTGGACTGTCTCTCAACACAGAGGAGAAATATATTGGAGAAAGAG 2903  
Db 2866 GCAGCTGGGCTGTGGCTGGCCATGTGACCCCGAGAAATGCCAGTTTGGCCAGGCTC 2925  
QY 2904 TGTTCGTGTGGGGACACAGGTTTCAATTTCCATGCTTCGCAATGTATCTGACCCATATT 2963  
Db 2926 AGGACCCATTTCTCTGGATGTGCTGCTCAGGACACGAGTCTTACCTGTGGAGCTG 2985  
QY 2964 TCAATGACAGTTCTTGGAGCACCTCCCTGTATCTCATGAAATATCTGTCTGTGATCTG 3023  
Db 2986 CCCCACAAATGGCTGGCTCTCCACAACTGTGGCCATCATGAGATGCTGTGCTCATCTG 3045  
QY 3024 CACAGAAAGCTTACCCAGCCACTGTTTCCATGCTTCGCAATGTATCTGACCCATATT 3083  
Db 3046 CTCAGCTGTCTCAGTCCCAGTCAACG---CCAGGCCAGATACTTGGCTGACCACTT 3102  
QY 3084 GTCTCAGTTCCAGAGGCGAGTGTGCTGTGCTTAGAGGACAAACGGCTCCGCTAGT 3143  
Db 3103 ACCGCTGATGACATGAGTATCTGAATCCAGTT-----TGGCTCTGAGGCTGT 3150  
QY 3144 GGATGGGACAGCCGCTGTGCGGGAGAGTAGAGATCTATCAGCAGCGCTTCTGGGGCAC 3203  
Db 3151 GAATGGAGTGCAGAGTGTGAGGCCGAGTGGAGTCTCTGTATCAGGCTCTCTGGGAAAC 3210  
QY 3204 CATCTGTATGACGGCTGGACCTGAGCGATGCCACGTGTGTGTCAAAAGCTGGGCTG 3263  
Db 3211 CGTGTGTATGACAGTGGGACACCAATGATGCCAATGTGTCTGTCAGGACGCTGGGCTG 3270  
QY 3264 TGGAGTGGCTTCAATGCCACGCTCTCTGCTCAGTTTGGGAGGCTCAGGCGCCATCTG 3323  
Db 3271 TGGCTGGGCGATGTGCGCCCGAGGAATGCCGGTTTGGCCAGGCTCAGGACCACTTGT 3330  
QY 3324 GCTGATGACCTGAACTGCACAGGAAAGGAGTCCCACTTGTGCGAGTGCCTTCCCGGG 3383  
Db 3331 CCTGGATGATGTGCGCTGTCTCAGGGAATGAGTCTACCTGTGGAGCTGCCCCACAAAG 3390  
QY 3384 CTGGGGGACAGCAGTCTCAGGACACAGGAGGAGGCTCAGGCTCATCTGCTCAG 3436  
Db 3391 CTGGCTCACCACAACTGTGGCCATCATCAAGAGCGCTGGTGTCTGCTCAG 3443

## RESULT 2

US-08-470-350B-1  
; Sequence 1, Application US/08470350B  
; Patent No. 5684126  
; GENERAL INFORMATION:  
; APPLICANT: Li, Xiao  
; APPLICANT: Solomon H  
; TITLE OF INVENTION: Ebnerin: A Secreted von Ebner's Gland  
; TITLE OF INVENTION: Protein Associated with Taste Buds  
; NUMBER OF INVENTIONS: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,350B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Wolfe, Susan A
REGISTRATION NUMBER: 33,568
REFERENCE/DOCKET NUMBER: 01107,48790
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4360 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3963
US-08-470-350B-1

Query Match
Best Local Similarity 3.2%; Score 149.6; DB 1; Length 4360;
Matches 407; Conservative 0; Mismatches 324; Indels 33; Gaps 3;

QY 2722 GATCTCGAGTGTGGTGGTCCAGTCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTT 2781
DB 523 GCTGTGAGGCTGGTGAATGAGGAGACAGGTGTGGGTCTGTGGAGATCTTTACCAG 582

QY 2782 GGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATGCCCGTGTCTATGC 2841
DB 583 GGTTCCTGGGACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAAGTGTGTGC 642

QY 2842 AGACAGTCACTGTGGAGTGTCTCTCAACACAGGAGAAATATATTGGAAAGA 2901
DB 643 AGCAGTGGGCTGTGGCTGGGCTTCTGCCCCAGGAAGTGGCCAGTTTGGACAGGC 702

QY 2902 AGTCTGTGTGTGGGACACAGTTCATTGCTTAGGGAATAGTCACTTCTCGATAAC 2961
DB 703 TCTGGTCCATGTTCTGGATGAGCTGGCTGTAGAGACATGAGGCTATCTGTGGAGC 762

QY 2962 TGTCAAAATGACAGTCTTGTGGAGCACCTCCCTGTATCCATGGAATACTGTCTGTGATC 3021
DB 763 TGTCTCCACCGAGCTGCTCTCTCATAACTGTGACATCAGGAGGATGCTGGAGTCATC 822

QY 3022 TGCAC-----AGAACCTTGACCCAGCCACTGTTTCCATGCCCTGCGAAATGT 3068
DB 823 TGTTCATATTCTCAAAAGACAGTCCACACCCGATTTCTAAACAAAGCAGTCCACACC 882

QY 3069 ATCTGACCCATATTGTCTGAGTTCAGAGGAGGAGTCT-TTATCTGCTTAGAGACA 3127
DB 883 GGTGTGGAAACCCGGGGGCAAAATAACGATGTGTCTATGAGCCGCAACAGACCACA 942

QY 3128 AACG-----GCTCCGCTAGTGGGACAGCGCGTGTGCCGGG 3168
DB 943 GACGCAACAGATTCTGTTGGTGTGAGGCTGTGATGAGGAGACAGGTGTGAGGT 1002

QY 3169 AGATAGAGATCTATACAGAGGCTTCTGGGGCACCATCTGTGTGATGACGCTGGGACCTG 3228
DB 1003 CGTGTGAGATCTTTACAGGGTTCCTGGGGTACCGCTGTGTGACGACAGCTGGACACC 1062

QY 3229 AGCATGCCACAGTGTGTCAAAAGCTGGGCTGTGAGTGGCTTCAATGCCACGGTC 3288
DB 1063 AAGGATCCCAACGTGGTGTGAGGACAGTGTGTGCTGGGCTTGTCTGCCCCAGGA 1122

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/341,587
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mollenhauer, Jan
REGISTRATION NUMBER: 33,568
REFERENCE/DOCKET NUMBER: 01107,48790
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4360 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3963
US-08-470-350B-1

Query Match
Best Local Similarity 3.2%; Score 149.6; DB 1; Length 4360;
Matches 407; Conservative 0; Mismatches 324; Indels 33; Gaps 3;

QY 2722 GATCTCGAGTGTGGTGGTCCAGTCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTT 2781
DB 523 GCTGTGAGGCTGGTGAATGAGGAGACAGGTGTGGGTCTGTGGAGATCTTTACCAG 582

QY 2782 GGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATGCCCGTGTCTATGC 2841
DB 583 GGTTCCTGGGACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAAGTGTGTGC 642

QY 2842 AGACAGTCACTGTGGAGTGTCTCTCAACACAGGAGAAATATATTGGAAAGA 2901
DB 643 AGCAGTGGGCTGTGGCTGGGCTTCTGCCCCAGGAAGTGGCCAGTTTGGACAGGC 702

QY 2902 AGTCTGTGTGTGGGACACAGTTCATTGCTTAGGGAATAGTCACTTCTCGATAAC 2961
DB 703 TCTGGTCCATGTTCTGGATGAGCTGGCTGTAGAGACATGAGGCTATCTGTGGAGC 762

QY 2962 TGTCAAAATGACAGTCTTGTGGAGCACCTCCCTGTATCCATGGAATACTGTCTGTGATC 3021
DB 763 TGTCTCCACCGAGCTGCTCTCTCATAACTGTGACATCAGGAGGATGCTGGAGTCATC 822

QY 3022 TGCAC-----AGAACCTTGACCCAGCCACTGTTTCCATGCCCTGCGAAATGT 3068
DB 823 TGTTCATATTCTCAAAAGACAGTCCACACCCGATTTCTAAACAAAGCAGTCCACACC 882

QY 3069 ATCTGACCCATATTGTCTGAGTTCAGAGGAGGAGTCT-TTATCTGCTTAGAGACA 3127
DB 883 GGTGTGGAAACCCGGGGGCAAAATAACGATGTGTCTATGAGCCGCAACAGACCACA 942

QY 3128 AACG-----GCTCCGCTAGTGGGACAGCGCGTGTGCCGGG 3168
DB 943 GACGCAACAGATTCTGTTGGTGTGAGGCTGTGATGAGGAGACAGGTGTGAGGT 1002

QY 3169 AGATAGAGATCTATACAGAGGCTTCTGGGGCACCATCTGTGTGATGACGCTGGGACCTG 3228
DB 1003 CGTGTGAGATCTTTACAGGGTTCCTGGGGTACCGCTGTGTGACGACAGCTGGACACC 1062

QY 3229 AGCATGCCACAGTGTGTCAAAAGCTGGGCTGTGAGTGGCTTCAATGCCACGGTC 3288
DB 1063 AAGGATCCCAACGTGGTGTGAGGACAGTGTGTGCTGGGCTTGTCTGCCCCAGGA 1122

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/341,587
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mollenhauer, Jan
REGISTRATION NUMBER: 33,568
REFERENCE/DOCKET NUMBER: 01107,48790
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4360 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
FEATURE:
NAME/KEY: CDS
LOCATION: 94..3963
US-08-470-350B-1

Query Match
Best Local Similarity 3.2%; Score 149.6; DB 1; Length 4360;
Matches 407; Conservative 0; Mismatches 324; Indels 33; Gaps 3;

QY 2722 GATCTCGAGTGTGGTGGTCCAGTCCAGTGTGACGGGCAAGTGGAGATCAACGTGCTT 2781
DB 523 GCTGTGAGGCTGGTGAATGAGGAGACAGGTGTGGGTCTGTGGAGATCTTTACCAG 582

QY 2782 GGACACTGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATGCCCGTGTCTATGC 2841
DB 583 GGTTCCTGGGACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAAGTGTGTGC 642

QY 2842 AGACAGTCACTGTGGAGTGTCTCTCAACACAGGAGAAATATATTGGAAAGA 2901
DB 643 AGCAGTGGGCTGTGGCTGGGCTTCTGCCCCAGGAAGTGGCCAGTTTGGACAGGC 702

QY 2902 AGTCTGTGTGTGGGACACAGTTCATTGCTTAGGGAATAGTCACTTCTCGATAAC 2961
DB 703 TCTGGTCCATGTTCTGGATGAGCTGGCTGTAGAGACATGAGGCTATCTGTGGAGC 762

QY 2962 TGTCAAAATGACAGTCTTGTGGAGCACCTCCCTGTATCCATGGAATACTGTCTGTGATC 3021
DB 763 TGTCTCCACCGAGCTGCTCTCTCATAACTGTGACATCAGGAGGATGCTGGAGTCATC 822

QY 3022 TGCAC-----AGAACCTTGACCCAGCCACTGTTTCCATGCCCTGCGAAATGT 3068
DB 823 TGTTCATATTCTCAAAAGACAGTCCACACCCGATTTCTAAACAAAGCAGTCCACACC 882

QY 3069 ATCTGACCCATATTGTCTGAGTTCAGAGGAGGAGTCT-TTATCTGCTTAGAGACA 3127
DB 883 GGTGTGGAAACCCGGGGGCAAAATAACGATGTGTCTATGAGCCGCAACAGACCACA 942

QY 3128 AACG-----GCTCCGCTAGTGGGACAGCGCGTGTGCCGGG 3168
DB 943 GACGCAACAGATTCTGTTGGTGTGAGGCTGTGATGAGGAGACAGGTGTGAGGT 1002

QY 3169 AGATAGAGATCTATACAGAGGCTTCTGGGGCACCATCTGTGTGATGACGCTGGGACCTG 3228
DB 1003 CGTGTGAGATCTTTACAGGGTTCCTGGGGTACCGCTGTGTGACGACAGCTGGACACC 1062

QY 3229 AGCATGCCACAGTGTGTCAAAAGCTGGGCTGTGAGTGGCTTCAATGCCACGGTC 3288
DB 1063 AAGGATCCCAACGTGGTGTGAGGACAGTGTGTGCTGGGCTTGTCTGCCCCAGGA 1122
```

EARLIER APPLICATION NUMBER: PCT/DE98/00096  
EARLIER FILING DATE: 1998-01-09  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 28720  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-341-587-7

Query Match 3.0%; Score 139.2; DB 4; Length 28720;  
Best Local Similarity 66.1%; Pred. No. 3.4e-28;  
Matches 201; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 3133 CTCGCGCTAGTGGGACAGCCGCTGTGCGGGAGAGTAGAGATCTATCAGCAGCGC 3192  
DB 2944 CTGAGACTGGTGAATGGAGGTGACAGGTGTGAGGCCGAGTGGAGTCTCTATACCAAGC 3003  
QY 3193 TTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAA 3252  
DB 3004 TCCTGGGGCACCCTGTGTGATGACTACTGGGACACCAATGATGCCAACGTGCTGCAGG 3063  
QY 3253 AAGCTGGGCTGGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGTCA 3312  
DB 3064 CAGCTGGGCTGTGGCTGGGCCATGTACGCCCCAGGAATGCCAGTTTGGCCAGGGCTCA 3123  
QY 3313 GGGCCCATCTGGCTGATGACCTGAACCTGCACAGGAACGGAGTCCACCTGTGGCAGTGC 3372  
DB 3124 GAACCATTTGCTCTGGATGTGGCTGCTCAGACAGAGTCTTACCTGTGGAGCTGC 3183  
QY 3373 CTTTCCCGCGGTGGGGCAGCAGCTGCAGGACAAAGAGGAGCGCAGGGGTCTATCTGC 3432  
DB 3184 CCCCACAATGGCTGCTCCCAACAACCTGTGCCATCATGAAGATGCTGTGCTCATCTGC 3243  
QY 3433 TCAG 3436  
DB 3244 TCAG 3247

RESULT 5  
US-09-341-587-2  
Sequence 2, Application US/09341587  
Patent No. 6346606  
GENERAL INFORMATION:  
APPLICANT: Mollenhauer, Jan  
TITLE OF INVENTION: Protein Containing an SRCR Domain  
FILE REFERENCE: 4121-108  
CURRENT APPLICATION NUMBER: US/09/341,587  
CURRENT FILING DATE: 1999-08-31  
EARLIER APPLICATION NUMBER: PCT/DE98/00096  
EARLIER FILING DATE: 1998-01-09  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2001  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-341-587-2

Query Match 2.9%; Score 136; DB 4; Length 2001;  
Best Local Similarity 65.5%; Pred. No. 5.1e-28;  
Matches 199; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 3133 CTCGCGCTAGTGGGACAGCCGCTGTGCGGGAGAGTAGAGATCTATCAGCAGCGC 3192  
DB 184 CTGAGCTGGTGAATGGAGGTGACAGGTGTGAGGCCGAGTGGAGTCTCTATCAGAGC 243  
QY 3193 TTCTGGGGCACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAA 3252  
DB 244 TCCTGGGAACCCGTGTGTGATGACAGCTGGGACACCAATGATGCCAATGTGCTCTGCAGG 303  
QY 3253 AAGCTGGGCTGGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTTGGGAGGGTCA 3312

DB 304 CAGCTGGGCTGTGGCTGGGCCATGTGCGCCCAAGGAAATGCCGGTTTGGCCAGGGCTCA 363  
QY 3313 GGGCCCATCTGGCTGATGACCTGAACCTGCACAGAACGAGTCCACACTTGTGGCAGTGC 3372  
DB 364 GGACCATTTGCTCTGGATGATGTGGCTGCTCAGGAATGAGTCTACTCTGTGGAGCTGC 423  
QY 3373 CTTTCCCGCGGTGGGGCAGCAGCTGCAGGACCAAGAGGAGCGCAGGGGTCTATCTGC 3432  
DB 424 CCCCACAAGGCTGGCTCACCCACAACCTGTGCCATCATGAAGAGCGCTGTGCTCATCTGC 483  
QY 3433 TCAG 3436  
DB 484 TCAG 487

RESULT 6  
US-09-034-916-1  
Sequence 1, Application US/09034916  
Patent No. 6046314  
GENERAL INFORMATION:  
APPLICANT: GEBE, JOHN A.  
APPLICANT: SIADAK, ANTHONY W.  
APPLICANT: ARUFFO, ALEJANDRO A.  
TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR  
TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROBINS & ASSOCIATES  
STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
CITY: MENLO PARK  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94025

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/034,916  
FILING DATE: 04-MAR-1998  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/039,956  
FILING DATE: 06-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 5998-0020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 325-7812  
TELEFAX: (650) 325-7823  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2178 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 64...1101  
US-09-034-916-1

Query Match 2.8%; Score 127.6; DB 3; Length 2178;  
Best Local Similarity 51.1%; Pred. No. 1.3e-25;  
Matches 499; Conservative 0; Mismatches 439; Indels 38; Gaps 7;  
QY 3134 TCCTGCTAGTGGATGGGACAGCCGCTGTGCGGGAGAGTAGAGATCTATCAGCAGCGCT 3193  
DB 131 TCGCGCTGTGGGGGCTCCACCGCTGTGAAGGGCGGCTGGAGGTGGAACAAAGGCC 190

Qy 3194 TCTGGGCGACCATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAA 3253  
|||||  
Db 191 AGTGGGCGACCGTGTGTGATGACGGCTGGGACCTAAGGACGCTGGCTGTGTGTGCCGG 250  
Qy 3254 AGCTGGGCTGTGGAGTGGCCCTTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGGTGAG 3313  
|||||  
Db 251 AGCTGGGCTGTGGAGTGGCCCGGAGGACCCCTAGTGGTATTTGTATGAGCGACCGAG 310  
Qy 3314 GGCCCATCTGGCTGATGACCT-GAATGTCACAGGAGGAGTCCCACTTGTGGCAGTGC 3372  
|||||  
Db 311 AAAAGAGCAAAAGTCTCTATCAATCAGTCAGTTGCAGAGGAACAGATATGTTG 370  
Qy 3373 CTTTCCCGGCTGGGGGAG--CACGACTGCAGCAAGAGGAGGAGCGAGGGTCACT 3430  
|||||  
Db 371 CTCAGTGTGAGCAAGAAGTGTATGATTTTACATGATGAAGATGCTGGG-CATCG 429  
Qy 3431 GCTCAGAAATTCACAGCCTTGGGCTTACAGTGAAGTGAATGAA----- 3472  
|||||  
Db 430 TGTGAGAACCCAGAGAGCTCTTTCTCCCACTGCCAGAGGTTGTCAGGCTGGCTGACGGC 489  
Qy 3473 -CAGAGAGCTGTGCTGGGAGATTTGAAGTCTTCTATAACGGGACCTGGGGCAGCGTGGC 3531  
|||||  
Db 490 CTTGGGCAATTCAGGGGAGCGTGAAGTGAAGCAGACCACTGATACCTGTTGTC 549  
Qy 3532 AGGAGGAATATCACACAGCCATAGCAGGCAATTTGTGCGAGCAGCTGGGCTGTGGGAG 3591  
|||||  
Db 550 CAGACAGGCTGGAGCTCCGGGCGCAAGGTTGTGCGGCGAGCTGGGATGTGGGAGG 609  
Qy 3592 AATGGAGT---TGTGAGCTGCCCCCTTATCTAAGACAGGCTGTGTTTATGTGGGTG 3648  
|||||  
Db 610 GCTGTACTGACTCAAAAACGCTGCAACAGCATGCTATGGCCGAAACCCATCTGGTGG 669  
Qy 3649 GATGATTCAGTGTCTTAAAGGATATCTCCATATGCGAGTGTCTGTGCCCATAGG 3708  
|||||  
Db 670 AGCCAGATGCTATGCTCAGGAGGAGCAACCTTCAAGATTTGCCCTTCTGGGCTTGG 729  
Qy 3709 GAGCGAAGAAATCTCAGCCCGAGCAAGAGAGCCTGGATCACATGTGAAGA-----TAGA 3762  
|||||  
Db 730 GGAAGAGACACTGCAACCATGATGAAGACACGCTGGGTGCAATGTGAAGATCCCTTTGAC 789  
Qy 3763 ATAAGAGTGGTGGAGGAGACACCGAGTGTCTTGGGAGTGGAGATCTGGCAGCAGGC 3822  
|||||  
Db 790 TTGAGACTAGTAGGAGGAGCAACCTCTGCTTGGGCGACTGGAGGTGCTGCACAAAGGC 849  
Qy 3823 TCTGGGCGACAGTGTGATGACTCTCTGGGACCTTGGCGAGGCGGAAGTGTGTCTGAG 3882  
|||||  
Db 850 GTATGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGACCGAGGTATGTCAG 909  
Qy 3883 CAGCTGGGCTGTGGCTCTGCTGTGGCTGCCCTGAGGGAGC-----CTTGGTTTGGCCAG 3936  
|||||  
Db 910 CAACCTGGGCTGTGGAACTCCCTCTCTCTCCCTTCAGAGACCGGAAATGCTATGGCCT 969  
Qy 3937 GGAAGTGAACATCTGTTGGATGACATGCGGTGCAAGGAAATGATGATTTCTATGG 3996  
|||||  
Db 970 GGGGTGTGGCCCATCTGCTGTGATAATGTTGCTTGGGAGGAGGACGCTCCCTGGAG 1029  
Qy 3997 GACTGTACCGCAAAACCTTGGGAGCAGAGTGAATGTGGACACAAAGAAAGTGTGGCGTG 4056  
|||||  
Db 1030 CAGTCCAGACAGATTTTGGGGTTTCCAGACTCCACCCAGGAGAGATGTGCTGTC 1089  
Qy 4057 AGGTGCTCTGGACAT 4072  
|||||  
Db 1090 ATCTGCTCAGGATAGT 1105

## RESULT 7

US-08-477-674-9  
; Sequence 9, Application US/08477674  
; Patent No. 5644035  
; GENERAL INFORMATION:  
; APPLICANT: Roths, Kirston E.  
; APPLICANT: Halenbeck, Robert F.  
; APPLICANT: Taylor, Eric W.

APPLICANT: Wang, Alice M.  
APPLICANT: Caspitt, Clayton L.  
TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cetus Oncology Corporation  
STREET: 1400 Fifty-Third Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/477,674  
APPLICATION NUMBER: US/08/477,674  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,714  
FILING DATE:  
APPLICATION NUMBER: US/07/961,404  
FILING DATE: 15-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Kenneth M.  
REGISTRATION NUMBER: 34,174  
REFERENCE/DOCKET NUMBER: 2595.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 420-3152  
TELEFAX: (510) 658-5470  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2285 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-477-674-9

Query Match 2.6%; Score 118.6; DB 1; Length 2285;  
Best Local Similarity 62.1%; Pred. No. 4.4e-23;  
Matches 187; Conservative 0; Mismatches 114; Indels 0; Gaps 0;  
Qy 3134 TCCGCTAGTGGATGGGACACGCCCTGTGCGGGAGAGTAGAGATCTATCAGAGCGCT 3193  
|||  
Db 250 TCGGCTGGCCGATGGGGCGCCACCAACGAGCGCGGTGGAGATCTTCTACAGAGGCC 309  
Qy 3194 TCTGGGCGACCATCTGTGATGACGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAA 3253  
|||  
Db 310 AGTGGGCGACTGTGTGACAACTGTGGGACCTGACTGATGCCAGCGCTGCTGCCGG 369  
Qy 3254 AGCTGGGCTGTGGAGTGGCCCTTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGGTGAG 3313  
|||  
Db 370 CCTGGGCTTCAGAAACGCCACCGAGCTCTGGGAGAGCTGCTTGGGCAAGGATGAG 429  
Qy 3314 GGCCCATCTGGCTGATGACCTGAAGTGCACAGGAACGAGTCCCACTTGTGGCAGTGC 3373  
|||  
Db 430 CCCCCTATGCTGGAGGAGTCCAGTGCAGGGAACCGAGGCTCAGTCCGCGGCTGCA 489  
Qy 3374 CTTCCCGGCTGGGGGAGCAGCAGTGCAGGCAAGGAGGAGGAGGCTCATCTGCT 3433  
|||  
Db 490 AGTCCCTGGGCTGCTGAAGACCACTGCAGGCAAGAGAGAGCGCTGGTGTGCTGCA 549  
Qy 3434 C 3434  
Db 550 C 550  
RESULT 8  
US-08-473-791-9

[illegible]

	Query Match	2.6%	Score 118.6;	DB 2;	Length 2285;
	Best Local Similarity	62.1%;	Pred. No. 4.4e-23;		
	Matches 187;	Conservative	0;	Mismatches 114;	Indels 0; Gaps 0;
QY	3134	TCGCGCTAGTGGATGGGGACAGCGCGTGTGCCGGAGAGTAGAGATCTATCACGACGGCT	3193		
DB	250	TGCGGCTTGGCGATGGGGGCGCCACCAACGAGGCCCGGTGAGATCTTCTACAGAGGCC	309		
QY	3194	TCTGGGGCACCATCTTGATGACGGCTGGGACCTGAGCGATGCCCACTGGTGTGTCAAA	3253		
DB	310	AGTGGGCACTGTGTGTGACAACCTGTGGGACCTGACTGATGCCAGCGCTCTGTGCCGGG	369		
QY	3254	AGCTGGGCTGTGGAGTGGCGCTTCAATGCCAGCGTCTCTGTCTCACTTTTGGGGAGGGGTCA	3313		
DB	370	CCCTGGGCTTCGAAACGCCACCCAGGCTCTGGGCAGAGCTGCCTTCGGGCAAGGATCAG	429		
QY	3314	GGCCCATCTGGCTGGATGACCTGAACTGCACAGAACGGAGTCCCACTTGTGGCAGTGCC	3373		
DB	430	GCCCCATCATCTGGACGAGGTCCAGTGTGACGGGAACCGGCCCTCACTGGCCGACTGCA	489		
QY	3374	CTTCCCGCGGCTGGGGGCGACGACGACTTCAGGCGACAAAGGAGGACGCGAGGGGTCACTCTG	3433		



; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2262  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-448-076-3

Query Match 2.3%; Score 108.2; DB 4; Length 2262;  
Best Local Similarity 61.6%; Pred. No. 3.7e-20;  
Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 3157 CGCTGCGCGGAGAGTAGATATACAGAGCGGCTTCTGGGGACCACTCTGTGATGAC 3216  
DB 157 CCTACGAGGCGCGGTGGAGATACAGGAGCTGTGAATGGGACCACTCTGCGATGAT 216  
QY 3217 GCGTGGGACCTGAGGATGCCACAGTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCCTTC 3276  
DB 217 GACTTTCACGCTGAGGCTGCCACATCTCTCCGGGAGCTGGCTTCACAGAGGCCACA 276  
QY 3277 AATGCCACGGTCTCTGCTCACTTCTGGGAGGGGTACAGGCCCATCTGGCTGGATGACCTG 3336  
DB 277 GCGTGGACCCACAGTGCCTTGTGGAGGCTTCCCGGCTGGAGGCTGGCTGGACAACTTG 336  
QY 3337 AACTGCACAGGAGGAGTCCCACTTGTGGAGTGCCTTCCCGGCTGGGAGGAGCAC 3396  
DB 337 AGCTGCAGTGGACCGAGGAGTGTGACTGAATGTGCTCCCGGCTGGGGAACAGT 396  
QY 3397 GACTGCAGGACCAAGGAGGAGGAGGCTCATCTGCTCAGA 3437  
DB 397 GACTGTACGCACGATGAGGATGCTGGGGTCACTCTGCAAGA 437

## RESULT 13

US-09-702-572-3  
; Sequence 3, Application US/09702572  
; Patent No. 6391602

; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND  
; FILE OF INVENTION: USES THEREFOR  
; FILE REFERENCE: MNI-073  
; CURRENT APPLICATION NUMBER: US/09/702,572  
; PRIOR FILING DATE: 2000-10-31  
; PRIOR APPLICATION NUMBER: 09/276,400  
; PRIOR FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2262  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-702-572-3

Query Match 2.3%; Score 108.2; DB 4; Length 2262;  
Best Local Similarity 61.6%; Pred. No. 3.7e-20;  
Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 3157 CGCTGCGCGGAGAGTAGATATACAGAGCGGCTTCTGGGGACCACTCTGTGATGAC 3216  
DB 157 CCTACGAGGCGCGGTGGAGATACAGGAGCTGTGAATGGGACCACTCTGCGATGAT 216  
QY 3217 GCGTGGGACCTGAGGATGCCACAGTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCCTTC 3276  
DB 217 GACTTTCACGCTGAGGCTGCCACATCTCTCCGGGAGCTGGCTTCACAGAGGCCACA 276  
QY 3277 AATGCCACGGTCTCTGCTCACTTCTGGGAGGGGTACAGGCCCATCTGGCTGGATGACCTG 3336  
DB 277 GCGTGGACCCACAGTGCCTTGTGGAGGCTTCCCGGCTGGAGGCTGGCTGGACAACTTG 336  
QY 3337 AACTGCACAGGAGGAGTCCCACTTGTGGAGTGCCTTCCCGGCTGGGAGGAGCAC 3396  
DB 337 AGCTGCAGTGGACCGAGGAGTGTGACTGAATGTGCTCCCGGCTGGGGAACAGT 396

QY 3397 GACTGCAGGACCAAGGAGGAGGAGGCTCATCTGCTCAGA 3437  
DB 397 GACTGTACGCACGATGAGGATGCTGGGGTCACTCTGCAAGA 437

## RESULT 14

US-09-276-400-1

; Sequence 1, Application US/09276400  
; Patent No. 6140056

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran  
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND  
; FILE OF INVENTION: USES THEREFOR  
; FILE REFERENCE: MNI-073  
; CURRENT APPLICATION NUMBER: US/09/276,400  
; CURRENT FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2920  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (143)..(2401)  
US-09-276-400-1

Query Match 2.3%; Score 108.2; DB 3; Length 2920;  
Best Local Similarity 61.6%; Pred. No. 4.3e-20;

Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 3157 CGCTGTCGCGGAGAGTAGATATACAGAGCGGCTTCTGGGGACCACTCTGTGATGAC 3216  
DB 299 CCTACGAGGCGCGGTGGAGATACAGGAGCTGTGAATGGGACCACTCTGCGATGAT 358  
QY 3217 GCGTGGGACCTGAGGATGCCACAGTGTGTCTCAAAAGCTGGGCTGTGGAGTGGCCCTTC 3276  
DB 359 GACTTTCACGCTGAGGCTGCCACATCTCTCCGGGAGCTGGCTTCACAGAGGCCACA 418  
QY 3277 AATGCCACGGTCTCTGCTCACTTCTGGGAGGGGTACAGGCCCATCTGGCTGGATGACCTG 3336  
DB 419 GCGTGGACCCACAGTGCCTTGTGGAGGAGGCTCATCTGCTCAGA 3437  
QY 3337 AACTGCACAGGAGGAGTCCCACTTGTGGAGTGCCTTCCCGGCTGGGAGGAGCAC 3396  
DB 479 AGCTGCAGTGGGAGGAGGAGTGTGACTGAATGTGCTCCCGGCTGGGGAACAGT 538  
QY 3397 GACTGCAGGACCAAGGAGGAGGAGGCTCATCTGCTCAGA 3437  
DB 539 GACTGTACGCACGATGAGGATGCTGGGGTCACTCTGCAAGA 579

## RESULT 15

US-09-448-076-1

; Sequence 1, Application US/09448076  
; Patent No. 6300092

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran et al.  
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN  
; FILE REFERENCE: MNI-073CP  
; CURRENT APPLICATION NUMBER: US/09/448,076  
; CURRENT FILING DATE: 1999-11-23  
; EARLIER APPLICATION NUMBER: 60/117,580  
; EARLIER FILING DATE: 1999-01-27  
; EARLIER APPLICATION NUMBER: 09/276,400  
; EARLIER FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2920  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: CDS  
; LOCATION: (143)...(2401)  
US-09-448-076-1

Query Match 2.38; Score 108.2; DB 4; Length 2920;  
Best Local Similarity 61.6%; Pred. No. 4.3e-20;  
Matches 173; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy	3157	CGCTGTGCCGGGAGAGTAGATCTATCACGACGGCTTCTGGGGCACCACATCTGTGATGAC	3216
Db	299	CCCTACGAGGCCCGGTGGAGATACAGCGAGCTGGTGAATGGGCACCACATCTGCGATGAT	358
Qy	3217	GGCTGGGACCTGAGCGATGCCACGTGGTGTCAAAAGCTGGGCTGTGGAGTGGCCCTTC	3276
Db	359	GACTTCACGCTGCAGGCTGCCACATCTCTGCCGGGAGCTGGGCTTCACAGAGGCCACA	418
Qy	3277	AATGCCACGGTCTCTGCTCACTTTGGGGAGGGTCCAGGCCCATCTGGCTGGATGACCTG	3336
Db	419	GGCTGGACCCACAGTGCCAAATATGGCCCTGGACAGCCCGCATCTGGCTGGCAACTTG	478
Qy	3337	AACTGCACAGGAAGGAGTCCCACTTGTGGCAGTGCCCTTCCCGCGGCTGGGGGCAGCAC	3396
Db	479	AGCTGCAGTGGGACCGACGAGTGTGACTGAATGTGCTCCCGGGGCTGGGGGAACAGT	538
Qy	3397	GACTGCAGGCACAGGAGGACGACGAGGGGTATCTGCTCAGA	3437
Db	539	GACTGTACGCACGATGAGGATGCTGGGGTCTATCTGCAAGA	579

Search completed: May 12, 2003, 06:29:49  
Job time : 374 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2003, 02:05:59 ; Search time 608 Seconds  
(without alignments)  
17141.856 Million cell updates/sec

Title: US-09-759-130B-379

Perfect score: 4628

Sequence: 1 gcggcgctgcgactagatga.....aaacggcgcggtgggtgcac 4628

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4628	100.0	4628	22	AAF45123 Human TANGO 234 CD
2	4359	94.2	4359	22	AAF45124 Human TANGO 234 OR
3	3459.8	74.8	3670	22	AAS01217 DNA encoding human
4	2761.2	59.7	2800	22	AAS44609 Human full-length
5	1448.2	31.3	3716	22	AAI58649 Human polynucleoti
6	1448.2	31.3	3834	22	AAH98740 Human EST-derived
7	1448.2	31.3	3834	22	AAI60435 Human polynucleoti
8	1448.2	31.3	3834	22	AAI60436 Human polynucleoti
9	1409.4	30.5	3707	22	AAS94922 Human DNA sequence

10	1339.2	28.9	3811	22	AAI58650 Human polynucleoti
11	886.8	19.2	4308	22	AAF45141 Bovine WC1 ORF. B
12	740	16.0	820	21	AAA43017 Human secreted exp
13	642	13.9	690	22	AAH99258 Human protein enco
14	563	12.2	608	20	AAH99272 EST clone CFI18.
15	551.4	11.9	562	22	AAH99326 Human protein enco
16	447.4	9.7	449	24	ABL87362 Human ovarian canc
17	404	8.7	450	22	ABA08369 Human M160 precurs
18	379.2	8.2	527	22	AA544781 Human contig polyn
19	335.4	7.2	415	22	ABA09428 Human Cys-rich sca
20	325.8	7.0	339	24	ABN24521 Human ORFX polynuc
21	319.2	6.9	2697	24	AAS17590 DNA encoding novel
22	213.6	4.6	599	20	AAV90288 EST clone DK64. H
23	213	4.6	599	20	AAV87660 EST clone DK399.
24	207.8	4.5	702	21	AA443661 Mouse secreted exp
25	199.2	4.3	461	22	AAK54544 Human haematologic
26	197.8	4.3	5802	19	AAV49652 Human SRCR protein
27	197.6	4.3	2562	20	AAI19024 Human serine prote
28	197.6	4.3	3350	20	AAV72589 Human neurotysin
29	193.2	4.2	2614	20	AAI19027 Mouse serine prote
30	191.6	4.1	2376	20	AAV72590 Mouse neurotysin
31	173.4	3.7	444	21	AAC78302 Human cancer assoc
32	167	3.6	474	22	AAH32999 Human colon cancer
33	162	3.5	1269	22	AAH32999 Human colon cancer
34	149.6	3.2	4344	24	ABK63805 Human sbg14862SPER
35	149.6	3.2	4360	18	AAT44068 Rat sequence diffe
36	143.8	3.1	1165	22	ABA08386 Human M130 antigen
37	143.6	3.1	1589	22	AAH33692 Human colon cancer
38	143.2	3.1	697	21	AAC99009 Human pancreatic c
39	140.8	3.0	29598	19	AAV49654 Human SC2 DNA. HO
40	140.2	3.0	1659	22	AAI16354 Human sbg14862SPER
41	139.2	3.0	28720	19	AAV49655 Human SC3 DNA. HO
42	136	2.9	2000	19	AAV49651 Human SRCR protein
43	134.4	2.9	713	20	AAV88878 EST clone HV435.
44	127.8	2.8	1591	19	AAV49556 Human liver cell c
45	127.2	2.7	584	22	AAH98116 Murine 7-transmemb

## ALIGNMENTS

RESULT 1

AAF45123

ID AAF45123 standard; cDNA; 4628 BP.

XX AAF45123;

AC AAF45123;

XX 30-MAR-2001 (first entry)

DT Human TANGO 234 cDNA.

DE Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;

DE central nervous system; focal brain disorder; bipolar affective disorder;

DE global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

DE senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

DE Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

DE neuropsychiatric; psychoactive substance use; anxiety; ss.

XX Homo sapiens.

XX WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.



Db 1741 ATTGTAACTCTCAGGTGATGCAACATGGGCCCTGAGGCTGGTGGCGGCAGCAACCGC 1800  
Qy 1801 TGCTCGGAAGACCTGGAGGTGCTACTTTCAGGACGGTGGGCACAGTGTGTGATGACGGC 1860  
Db 1801 TGCTCGGAAGACCTGGAGGTGCTACTTTCAGGACGGTGGGCACAGTGTGTGATGACGGC 1860  
Qy 1861 TGGAAACAGTAAAGTGCAGCTGTGGTGTGTAGCCAGCTGGGACTGCCCATCTTCTATCAT 1920  
Db 1861 TGGAAACAGTAAAGTGCAGCTGTGGTGTGTAGCCAGCTGGGACTGCCCATCTTCTATCAT 1920  
Qy 1921 GGCATGGTCTGGGAAAGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980  
Db 1921 GGCATGGTCTGGGAAAGCTTCTACAGGATATGGAAAAATTTGGCTCGATGATGTTTCC 1980  
Qy 1981 TGTGATGAGATGAGTCAAGTCTCTGGTCAATGAGCAACAGTGGTGGGAAATTAATGAC 2040  
Db 1981 TGTGATGAGATGAGTCAAGTCTCTGGTCAATGAGCAACAGTGGTGGGAAATTAATGAC 2040  
Qy 2041 TGCAGTCAAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGAGCTGAGG 2100  
Db 2041 TGCAGTCAAGTGAAGATGTTGGAGTGATCTGTTCTGATGCATCGGATATGAGCTGAGG 2100  
Qy 2101 CTTGTGGTGAAGCAGCAGTGTGCTGGGAAAGTTGAGGTGAATGCTCCAGGGTCCCGTG 2160  
Db 2101 CTTGTGGTGAAGCAGCAGTGTGCTGGGAAAGTTGAGGTGAATGCTCCAGGGTCCCGTG 2160  
Qy 2161 GGAATTCCTGCTGCTTAATGGCTGGGAAATGAACATGCTGAAGTTGTTGCAGGCAACTT 2220  
Db 2161 GGAATTCCTGCTGCTTAATGGCTGGGAAATGAACATGCTGAAGTTGTTGCAGGCAACTT 2220  
Qy 2221 GAATGTGGGCTGCAATCAGGCTCTCCAGAGAGCTCATTTTCAGAGAAAGCAATACAC 2280  
Db 2221 GAATGTGGGCTGCAATCAGGCTCTCCAGAGAGCTCATTTTCAGAGAAAGCAATACAC 2280  
Qy 2281 ATCTTAATGCTCAATTCGGCTGCACATGGAGGGAAGCCTCTCTCTGGGATGTATAGGA 2340  
Db 2281 ATCTTAATGCTCAATTCGGCTGCACATGGAGGGAAGCCTCTCTCTGGGATGTATAGGA 2340  
Qy 2341 TGGAGTGGAAACAGACATGCTGTGCTTAATATGGAAGCAAGTTTTCATCTGCTCAGCC 2400  
Db 2341 TGGAGTGGAAACAGACATGCTGTGCTTAATATGGAAGCAAGTTTTCATCTGCTCAGCC 2400  
Qy 2401 CACAGGACGCCAGCTGGTGGAGCTGATATGCCCTGCTTGGAGCTGTGAGAGTGAAG 2460  
Db 2401 CACAGGACGCCAGCTGGTGGAGCTGATATGCCCTGCTTGGAGCTGTGAGAGTGAAG 2460  
Qy 2461 CATCAGACACATGCGCTGCTGTGATGCTGATTTCTCTCTCATGCTGCCAATGTG 2520  
Db 2461 CATCAGACACATGCGCTGCTGTGATGCTGATTTCTCTCTCATGCTGCCAATGTG 2520  
Qy 2521 CTGTCCAGAGAAATTAATTTGGAGATGCCATATCTCTTTCTGTGGAGATCACTTTTGA 2580  
Db 2521 CTGTCCAGAGAAATTAATTTGGAGATGCCATATCTCTTTCTGTGGAGATCACTTTTGA 2580  
Qy 2581 AAAGGAAATGCTAACTTTGGGCGGAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTT 2640  
Db 2581 AAAGGAAATGCTAACTTTGGGCGGAAAGTTCCAGTGTGAAGGAGTGAACACTCACCTT 2640  
Qy 2641 GCATATGCCCATTTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAGATTGGA 2700  
Db 2641 GCATATGCCCATTTGTTCAACATCCGGAAGACACTTGTATCCACAGCAGAGAGATTGGA 2700  
Qy 2701 GTTGTCTCTCCGATATACAGATGCTCCGACTTGTGAATGCAATCCAGTGTGACGGG 2760  
Db 2701 GTTGTCTCTCCGATATACAGATGCTCCGACTTGTGAATGCAATCCAGTGTGACGGG 2760  
Qy 2761 CAAGTGGAGATCAAGCTGCTTGGACACTGGGCTCACTGTGTGACACCCACTGGGACCCA 2820  
Db 2761 CAAGTGGAGATCAAGCTGCTTGGACACTGGGCTCACTGTGTGACACCCACTGGGACCCA 2820  
Qy 2821 GAAGATGCCCTGTGTTATGACAGACAGCTCAGCTGTGGGACTGCTCTCAACCCACAGGA 2880  
Db 2821 GAAGATGCCCTGTGTTATGACAGACAGCTCAGCTGTGGGACTGCTCTCAACCCACAGGA 2880

Qy 2881 GGAAATATATTGGAGAAAGAGTTCGTGTGGGACACAGGTTTCATTGCTTAGGG 2940  
Db 2881 GGAAATATATTGGAGAAAGAGTTCGTGTGGGACACAGGTTTCATTGCTTAGGG 2940  
Qy 2941 AATGAGTCACCTTCGGATAACTGTCAAAATGACAGTTCCTGGAGCACTCCCTGTATCAT 3000  
Db 2941 AATGAGTCACCTTCGGATAACTGTCAAAATGACAGTTCCTGGAGCACTCCCTGTATCAT 3000  
Qy 3001 GGAATACTGCTCTGTGATCTGCACAGAAAGCCCTGACCCAGCCACTGTTTCCATGCGCTC 3060  
Db 3001 GGAATACTGCTCTGTGATCTGCACAGAAAGCCCTGACCCAGCCACTGTTTCCATGCGCTC 3060  
Qy 3061 GCAATATGATGACCCCATATTTGCTCAGTTCAGAGGGCAGTGTTCATCTGCTTTA 3120  
Db 3061 GCAATATGATGACCCCATATTTGCTCAGTTCAGAGGGCAGTGTTCATCTGCTTTA 3120  
Qy 3121 GAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTGTGCCGGGAGAGTAGATC 3180  
Db 3121 GAGGACAAACGGCTCCGCTAGTGGATGGGACAGCCGCTGTGCCGGGAGAGTAGATC 3180  
Qy 3181 TATCACGACGGCTTCTGGGACACCATCTGTGATGACGCTGGGACCTGAGCGATGCCAC 3240  
Db 3181 TATCACGACGGCTTCTGGGACACCATCTGTGATGACGCTGGGACCTGAGCGATGCCAC 3240  
Qy 3241 GTGGTGTCTCAAAAGCTGGGCTGTGGAGTGGCTTCAATGCCACGCTCTCTGCTCACTTT 3300  
Db 3241 GTGGTGTCTCAAAAGCTGGGCTGTGGAGTGGCTTCAATGCCACGCTCTCTGCTCACTTT 3300  
Qy 3301 GGGAGGGGTGAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGAACGAGGTCCAC 3360  
Db 3301 GGGAGGGGTGAGGGCCCATCTGGCTGGATGACCTGAACTGCACAGAACGAGGTCCAC 3360  
Qy 3361 TTGTGGCAGTCCCTTCCCGGGCTGGGGCAGCAGCTGCAGGCAACAGGAGACCA 3420  
Db 3361 TTGTGGCAGTCCCTTCCCGGGCTGGGGCAGCAGCTGCAGGCAACAGGAGACCA 3420  
Qy 3421 GGGGTCTCTCTCAGAAATTCACAGCTTACAGTGTACAGTGAACCTGAAACAGAGAGC 3480  
Db 3421 GGGGTCTCTCTCAGAAATTCACAGCTTACAGTGTACAGTGAACCTGAAACAGAGAGC 3480  
Qy 3481 TGTGCTGGGAGATTGGAAGTCTTCTAATACGGGACCTGGGGAGCTGGCAGGAGGAAC 3540  
Db 3481 TGTGCTGGGAGATTGGAAGTCTTCTAATACGGGACCTGGGGAGCTGGCAGGAGGAAC 3540  
Qy 3541 ATCACACAGCATAGCAGGATCTGTGAGCAGCTGGGCTGTGGGAGAAATGGAGTT 3600  
Db 3541 ATCACACAGCATAGCAGGATCTGTGAGCAGCTGGGCTGTGGGAGAAATGGAGTT 3600  
Qy 3601 CTCAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTTCATGTGGTGGATGACATTGAG 3660  
Db 3601 CTCAGCTCGCCCTTTATCTAAGACAGGCTCTGTTTTCATGTGGTGGATGACATTGAG 3660  
Qy 3661 TGTCTTAAAGCATATCTCCATATGGCAGTGCCTGTCTGCCCATGGGAGGAAGATC 3720  
Db 3661 TGTCTTAAAGCATATCTCCATATGGCAGTGCCTGTCTGCCCATGGGAGGAAGATC 3720  
Qy 3721 TCCAGCCAGCAGAGAGACCTTGGATCACTGTGAAGATAGAAATAGAGTGGTGGAGGA 3780  
Db 3721 TCCAGCCAGCAGAGAGACCTTGGATCACTGTGAAGATAGAAATAGAGTGGTGGAGGA 3780  
Qy 3781 GACACCGAGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTGGGGCACAGTGTGT 3840  
Db 3781 GACACCGAGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTGGGGCACAGTGTGT 3840  
Qy 3841 GATGACTCCTGGGACCTGGCGAGCGGAGTGTGTGTGACAGCTGGGCTGTGGCTCT 3900  
Db 3841 GATGACTCCTGGGACCTGGCGAGCGGAGTGTGTGTGACAGCTGGGCTGTGGCTCT 3900  
Qy 3901 GCTCTGGCTGCCCTGAGGAGCGCTTCGTTTGGCCAGGAACTGGAACCATCTGGTTGAT 3960  
Db 3901 GCTCTGGCTGCCCTGAGGAGCGCTTCGTTTGGCCAGGAACTGGAACCATCTGGTTGAT 3960

```
QY 3961 GACATCGGTGCAAGGAAATGAGTCAATTTCTATGGAGTGTACGCCCAAAACCTGGGGA 4020
Db 3961 GACATCGGTGCAAGGAAATGAGTCAATTTCTATGGAGTGTACGCCCAAAACCTGGGGA 4020
QY 4021 CAGAGTGACTGTGGACACAGGAAGATGCTGGCGTCAAGTGTCTGGACAGTCGCTGAAA 4080
Db 4021 CAGAGTGACTGTGGACACAGGAAGATGCTGGCGTCAAGTGTCTGGACAGTCGCTGAAA 4080
QY 4081 TCACCTGAATGCCCTCCCTCAGTCAATTTAGCACATTTATTTATCCAGTATCTTTGGGCTCCCTT 4140
Db 4081 TCACCTGAATGCCCTCCCTCAGTCAATTTAGCACATTTATTTATCCAGTATCTTTGGGCTCCCTT 4140
QY 4141 CTCCTGGTCTGTTTATTTCTATTTCTACGTCGTGGTGGAGTTCAGAAACAAAACATCTG 4200
Db 4141 CTCCTGGTCTGTTTATTTCTATTTCTACGTCGTGGTGGAGTTCAGAAACAAAACATCTG 4200
QY 4201 CCCTCAGAGTTTCAACCAAGAGGAGGGTCTCTCGAGGAGAAATTTATTCATGAGATG 4260
Db 4201 CCCTCAGAGTTTCAACCAAGAGGAGGGTCTCTCGAGGAGAAATTTATTCATGAGATG 4260
QY 4261 GAGACCTGCTCAAGAGAGAGGCCACATGGGACAAAGACCTCAGATGACACCCCAAC 4320
Db 4261 GAGACCTGCTCAAGAGAGAGGCCACATGGGACAAAGACCTCAGATGACACCCCAAC 4320
QY 4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTCTTCTCGCTCTGAAGCC 4380
Db 4321 CATGGTTGTGAAGATGCTAGCGACACATCGCTGTTGGGAGTCTTCTCGCTCTGAAGCC 4380
QY 4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
Db 4381 ACAAAATGACTTTAGACTTCCAGGGCTCACCAGATCAACCTCTAAATATCTTTGAAGGAG 4440
QY 4441 ACAACAACTTTTAAATGAATAAAGAGAGTCAAGTTGCCCTATGGAAACCTTTGTCCAAA 4500
Db 4441 ACAACAACTTTTAAATGAATAAAGAGAGTCAAGTTGCCCTATGGAAACCTTTGTCCAAA 4500
QY 4501 TACATTTCTTGAAACATAGGAGAACAGCTAAATTTGATAAGACTGGTGATAATAAAAT 4560
Db 4501 TACATTTCTTGAAACATAGGAGAACAGCTAAATTTGATAAGACTGGTGATAATAAAAT 4560
QY 4561 TGAATATGATATACACTGTATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4620
Db 4561 TGAATATGATATACACTGTATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4620
QY 4621 GGGTCGAC 4628
Db 4621 GGGTCGAC 4628

RESULT 2
AAF45124
ID AAF45124 standard; cdna; 4359 BP.
XX AC AAF45124;
XX DT 30-MAR-2001 (first entry)
XX DE Human TANGO 234 ORF.
XX KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
XX OS Homo sapiens.
XX PN WO200077239-A2.
XX PD 21-DEC-2000.
XX PF 24-MAY-2000; 2000WO-US14858.
```

```
XX 14-JUN-1999; 99US-0333159.
PR (MILL-) MILLENNIUM PHARM INC.
XX PA
XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX P-PSDB; AAB66037.
XX WPI; 2001-032313/04.
XX DR
XX PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
XX screening assays and diagnostic assays and for the treatment of
XX neurological diseases such as Alzheimer's, Parkinson's and Huntington's
XX disease.
XX PS Claim 1; Fig 2; 359pp; English.
XX CC The present invention relates to TANGO or INTERCEPT proteins and coding
XX sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
XX AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
XX sequences are useful for the treatment of neurological disorders such as
XX central nervous system (CNS) disorders, CNS-related disorders, focal
XX brain disorders, global-diffuse cerebral disorders and other
XX neurological and cerebrovascular disorders. The CNS disorders include
XX Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
XX lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
XX autonomic function disorders such as hypertension and sleep disorders,
XX neuropsychiatric disorders, psychoactive substance use disorders,
XX anxiety, and bipolar affective disorder.
XX SQ Sequence 4359 BP; 1060 A; 881 C; 1303 G; 1115 T; 0 other;

Query Match 94.2%; Score 4359; DB 22; Length 4359;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 ATGATGCTGCTCAAAACCTGCGGCAATATGATTTTGGAGATGCTGTCTCATCAGAAC 87
Db 1 ATGATGCTGCTCAAAACCTGCGGCAATATGATTTTGGAGATGCTGTCTCATCAGAAC 60
QY 88 CTTTCTCTGCTGTGTAACCTTGCATCCTGCTCCTCAATTCCTCTTCTCATCAGCAGT 147
Db 61 CTTTCTCTGCTGTGTAACCTTGCATCCTGCTCCTCAATTCCTCTTCTCATCAGCAGT 120
QY 148 TTTAATGGAACAGATTTGGAGTTGAGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 207
Db 121 TTTAATGGAACAGATTTGGAGTTGAGCTGGTCAATGGAGACGGTCCCTGCTCTGGGACA 180
QY 208 GTGGAGGTGAAATTCAGGACAGTGGGGACCTGTGTGTCATGATGGTGGAAACACTACT 267
Db 181 GTGGAGGTGAAATTCAGGACAGTGGGGACCTGTGTGTCATGATGGTGGAAACACTACT 240
QY 268 GCCTCAACTGCTGCTGCAAAACAGCTTGGATGCCATTTCTTTCCGCAATGTTTCGTTTT 327
Db 241 GCCTCAACTGCTGCTGCAAAACAGCTTGGATGCCATTTCTTTCCGCAATGTTTCGTTTT 300
QY 328 GGACAAAGCGGTGACTAGACATGGAATAATTTGCTTCATGATGTTTCTCTGTTATGAAAT 387
Db 301 GGACAAAGCGGTGACTAGACATGGAATAATTTGCTTCATGATGTTTCTCTGTTATGAAAT 360
QY 388 GAGTCAGTCTCTGCGGAATGTCACACCGGAAATGGGAGGCCATACCTTTATCATGGA 447
Db 361 GAGTCAGTCTCTGCGGAATGTCACACCGGAAATGGGAGGCCATACCTTTATCATGGA 420
QY 448 GAAGATGTTGGTGAACCTGTTATGTTGAGGCAACTCTGGGTTTGGAGCTAGTGGATGGA 507
Db 421 GAAGATGTTGGTGAACCTGTTATGTTGAGGCAACTCTGGGTTTGGAGCTAGTGGATGGA 480
QY 508 AACAACTCTCTGTTGAGGAGAGTGGAGGTGAAATTTCCAGAAAGGTGGGGACTATATGT 567
Db 481 AACAACTCTCTGTTGAGGAGAGTGGAGGTGAAATTTCCAGAAAGGTGGGGACTATATGT 540
QY 568 GATGATGGGTGGAACCTTGAATGCTGCTGCCGTGTTGTCAGGCACTAGGATGTCCTCATCT 627
```

Db 541 GATGATGGGTGGAACCTGAATACTGCTCCCTGGTGTGCAGGCAACTAGGATGTCCACTCT 600  
Qy 628 TCTTTTATTTCTCTGGAGTTGTTAATAGCCCTGCTGTATTGGCCCCCAATTGGCTGAT 687  
Db 601 TCTTTTATTTCTCTGGAGTTGTTAATAGCCCTGCTGTATTGGCCCCCAATTGGCTGAT 660  
Qy 688 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAAATGCAGACATCGTGGATGGGA 747  
Db 661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAAATGCAGACATCGTGGATGGGA 720  
Qy 748 AATCATGACTCGAGTCACAATGAGAGATGTACATTAACCTGTTATGATAGTAGTATCTT 807  
Db 721 AATCATGACTCGAGTCACAATGAGAGATGTACATTAACCTGTTATGATAGTAGTATCTT 780  
Qy 808 GAACCTAAGGCTGTAGGTGGAACTAACCCGCTGTATGGGGAGAGTAGAGCTGAATCCAA 867  
Db 781 GAACCTAAGGCTGTAGGTGGAACTAACCCGCTGTATGGGGAGAGTAGAGCTGAATCCAA 840  
Qy 868 GGAAGGTGGGGACCGTATGCCACCATAAGTGGAAACAATGCTGCAGCTGATGTCGTATGC 927  
Db 841 GGAAGGTGGGGACCGTATGCCACCATAAGTGGAAACAATGCTGCAGCTGATGTCGTATGC 900  
Qy 928 AAGCAGTTGGGATGGAACCGCACTTCACCTCGCTGGCTTGCCTTCATTTGCAGTCAGG 987  
Db 901 AAGCAGTTGGGATGGAACCGCACTTCACCTCGCTGGCTTGCCTTCATTTGCAGTCAGG 960  
Qy 988 TCTGATGTTGATGGCTGTAGGTGTCTCCCTGCTCCGCTGAATGAATCTTTCTTTGGGAC 1047  
Db 961 TCTGATGTTGATGGCTGTAGGTGTCTCCCTGCTCCGCTGAATGAATCTTTCTTTGGGAC 1020  
Qy 1048 TCCAGACATTCGGGAACCGTCAATTTTGACTGTCTTCAATCAAAAGCATGTCTGTGATC 1107  
Db 1021 TCCAGACATTCGGGAACCGTCAATTTTGACTGTCTTCAATCAAAAGCATGTCTGTGATC 1080  
Qy 1108 TGCTCAGATGAGCAGATTTGGAACCTGCGACTAGCAGATGGAAGTAACAATTTGTCAGG 1167  
Db 1081 TGCTCAGATGAGCAGATTTGGAACCTGCGACTAGCAGATGGAAGTAACAATTTGTCAGG 1140  
Qy 1168 AGAGTAGAGGTGAGAAATTCATGAACAGTGGTGGCAATATGTGACCAAGCTGGAAGAT 1227  
Db 1141 AGAGTAGAGGTGAGAAATTCATGAACAGTGGTGGCAATATGTGACCAAGCTGGAAGAT 1200  
Qy 1228 GAACAAGCCCTTGTGGTTTGTGAAGCAGCTAGGATGTCGGTTCAGCGCTTTTGGCAGTCT 1287  
Db 1201 GAACAAGCCCTTGTGGTTTGTGAAGCAGCTAGGATGTCGGTTCAGCGCTTTTGGCAGTCT 1260  
Qy 1288 CGTGCTAAACCTAGTAATGAAGCTAGACATATTGGATTAACAGCATATCTTTGCACTGG 1347  
Db 1261 CGTGCTAAACCTAGTAATGAAGCTAGACATATTGGATTAACAGCATATCTTTGCACTGG 1320  
Qy 1348 AATGAGTCAGCTCTCTGGGACTGCAATATGATGAAAGCAAGCAAGCAATGCTTCCGA 1407  
Db 1321 AATGAGTCAGCTCTCTGGGACTGCAATATGATGAAAGCAAGCAAGCAATGCTTCCGA 1380  
Qy 1408 AGATCAGATGCTGGAGTAATTTGTTCTGATAGGCAGATCTGGACCTAAGGCTTCTCGGG 1467  
Db 1381 AGATCAGATGCTGGAGTAATTTGTTCTGATAGGCAGATCTGGACCTAAGGCTTCTCGGG 1440  
Qy 1468 GCTATAGCCCTGTTATGGGAGATGGAGGTGAATACCAAGGAGATGGGGGACTGTG 1527  
Db 1441 GCTATAGCCCTGTTATGGGAGATGGAGGTGAATACCAAGGAGATGGGGGACTGTG 1500  
Qy 1528 TGTATGACAGATGAGGACACAGGAATGCAGCTGTGTGTGTAACAATTTGGGATGTGA 1587  
Db 1501 TGTATGACAGATGAGGACACAGGAATGCAGCTGTGTGTGTAACAATTTGGGATGTGA 1560  
Qy 1588 AAGCCTATGCATGTGTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTGGCTG 1647  
Db 1561 AAGCCTATGCATGTGTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTGGCTG 1620  
Qy 1648 AATGACGTTTCTGCAATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATG 1707

Db 1621 GATGACGTTTCTTCATTTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1680  
Qy 1708 GGAAGACATAATGTGTACACAGAGAGGATGTGATTGTAACTGCTCAGGTGATCAACA 1767  
Db 1681 GGAAGACATAATGTGTGTACACAGAGAGGATGTGATTGTAACTGCTCAGGTGATCAACA 1740  
Qy 1768 TGGGSCCTGAGGCTGTGTGGGGGGGAGCAACCGCTCTCGGGAAGACTGAGAGGTGACTTT 1827  
Db 1741 TGGGSCCTGAGGCTGTGTGGGGGGGAGCAACCGCTCTCGGGAAGACTGAGAGGTGACTTT 1800  
Qy 1828 CAAGCACGCTGGGGCACAGTGTGTGATGACGGCTGGAACTGAACAGTGCAGCTGTGGTG 1887  
Db 1801 CAAGCACGCTGGGGCACAGTGTGTGATGACGGCTGGAACTGAACAGTGCAGCTGTGGTG 1860  
Qy 1888 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGATGATGGGTCTGGGAACGCTTCTACA 1947  
Db 1861 TGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGATGATGGGTCTGGGAACGCTTCTACA 1920  
Qy 1948 GATATGGAATAATTTGGCTTCGATGATGTTCTCTGTATGGAGATGAGTCAAGATCTCTGG 2007  
Db 1921 GATATGGAATAATTTGGCTTCGATGATGTTCTCTGTATGGAGATGAGTCAAGATCTCTGG 1980  
Qy 2008 TCATGCAGAACAGTGGGTGGGGAATAATGACTGCAGTGCAGTGCAGATGTTGGAGTG 2067  
Db 1981 TCATGCAGAACAGTGGGTGGGGAATAATGACTGCAGTGCAGTGCAGATGTTGGAGTG 2040  
Qy 2068 ATCTGTTCTGATGATCGATATGAGCTGAGGCTTGTGGGTGGAAGCAGGCTGTGCT 2127  
Db 2041 ATCTGTTCTGATGATCGATATGAGCTGAGGCTTGTGGGTGGAAGCAGGCTGTGCT 2100  
Qy 2128 GGAAGAAGTTGAGGTGAATGTTCCAGGTCGCTGGGAATTCGTGTGCTTAATGGCTGGGA 2187  
Db 2101 GGAAGAAGTTGAGGTGAATGTTCCAGGTCGCTGGGAATTCGTGTGCTTAATGGCTGGGA 2160  
Qy 2188 ATGAACATTTGCTGAAGTTGTTTTCAGGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCC 2247  
Db 2161 ATGAACATTTGCTGAAGTTGTTTTCAGGCAACTTGAATGTGGGTCTGCAATCAGGCTCTCC 2220  
Qy 2248 AGAGAGCTCTATTTTACAGAAAGAACATTAACATCTTAAATGTCGAATTTCTGGCTGCACT 2307  
Db 2221 AGAGAGCTCTATTTTACAGAAAGAACATTAACATCTTAAATGTCGAATTTCTGGCTGCACT 2280  
Qy 2308 GGAGGGGAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTCGCTGTCTAT 2367  
Db 2281 GGAGGGGAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACTCGCTGTCTAT 2340  
Qy 2368 TTAATATATGGAAGCAAGTTTGTATCTCAGCCACAGCAGCCAGGCTGTTGAGCT 2427  
Db 2341 TTAATATATGGAAGCAAGTTTGTATCTCAGCCACAGCAGCCAGGCTGTTGAGCT 2400  
Qy 2428 GATATGCCCTGCTCTGGGAGCTGTTGAAGTGAACATGACAGACATGCGCTCTCTCTGT 2487  
Db 2401 GATATGCCCTGCTCTGGGAGCTGTTGAAGTGAACATGACAGACATGCGCTCTCTCTGT 2460  
Qy 2488 GATTCGTATTTCTCTCTTCATGCTGCCAATGTGCTGTGTCAGAGAAATTAATTTGTGGAGAT 2547  
Db 2461 GATTCGTATTTCTCTCTTCATGCTGCCAATGTGCTGTGTCAGAGAAATTAATTTGTGGAGAT 2520  
Qy 2548 CCCATATCTTTCTCTCTGGAGATCATTGGAAGGAAGTGTCTAACTTTGGGCCGAA 2607  
Db 2521 CCCATATCTTTCTCTGGAGATCATTGGAAGGAAGTGTCTAACTTTGGGCCGAA 2580  
Qy 2608 AAGTTCCAGTGTGAAGGAGTGAACCTCACCTTTGATATATGCCCATTTCTTCAACATCCG 2667  
Db 2581 AAGTTCCAGTGTGAAGGAGTGAACCTCACCTTTGATATATGCCCATTTCTTCAACATCCG 2640  
Qy 2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCGGATATACAGATGTC 2727  
Db 2641 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCGGATATACAGATGTC 2700  
Qy 2728 CGACTTGTGAATGGCAATCCAGTGTGACGGCAGTGGAGATCAACCTGCTTGGACAC 2787  
Db 2701 CGACTTGTGAATGGCAATCCAGTGTGACGGCAGTGGAGATCAACCTGCTTGGACAC 2760

QY 2788 TGGGGCTCACTGTGTGACACCCCACTGGGACCCAGAAAGATGCCCGTGTCTTATGACAGAC 2847  
Db 2761 TGGGGCTCACTGTGTGACACCCCACTGGGACCCAGAAAGATGCCCGTGTCTTATGACAGAC 2820  
QY 2848 CTCAGCTGTGGGACTGCTCTCTCAACACACAGGAGGAAAATATATGAGAAAGAGTGT 2907  
Db 2821 CTCAGCTGTGGGACTGCTCTCTCAACACACAGGAGGAAAATATATGAGAAAGAGTGT 2880  
QY 2908 CGTGTGTGGGACACAGGTTTCATTGCTTAGGAAATGAGTCACTTCTGTGATACACTGCA 2967  
Db 2881 CGTGTGTGGGACACAGGTTTCATTGCTTAGGAAATGAGTCACTTCTGTGATACACTGCA 2940  
QY 2968 ATCAGAGTCTTCTGGAGCACCTCCCTGATCATGGAATGAGTCACTTCTGTGATACACTGCA 3027  
Db 2941 ATCAGAGTCTTCTGGAGCACCTCCCTGATCATGGAATGAGTCACTTCTGTGATACACTGCA 3000  
QY 3028 GGAAGCCTGACCCAGCCACTGTGTTCCATGCTCGCAAAATGATCTGACCCATATTTGTCT 3087  
Db 3001 GGAAGCCTGACCCAGCCACTGTGTTCCATGCTCGCAAAATGATCTGACCCATATTTGTCT 3060  
QY 3088 GCAGTTCAGAGGGCAGTGTGATCTGCTTAGAGGACAAACGGCTCCGCCCTAGTGGAT 3147  
Db 3061 GCAGTTCAGAGGGCAGTGTGATCTGCTTAGAGGACAAACGGCTCCGCCCTAGTGGAT 3120  
QY 3148 GGGGACAGCCGCTGTGCCGGGAGAGTAGATCTATCACAGCCGCTTCTGGGGCACCATC 3207  
Db 3121 GGGGACAGCCGCTGTGCCGGGAGAGTAGATCTATCACAGCCGCTTCTGGGGCACCATC 3180  
QY 3208 TGTGATGACGGCTGGGACCTGAGCGATGCCCCAGTGGTGTCAAAAGCTGGGCTGTGGA 3267  
Db 3181 TGTGATGACGGCTGGGACCTGAGCGATGCCCCAGTGGTGTCAAAAGCTGGGCTGTGGA 3240  
QY 3268 GTGGCTTCAATGACAGGCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGCTG 3327  
Db 3241 GTGGCTTCAATGACAGGCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGCTG 3300  
QY 3328 GATGACCTGAACCTGCACAGGAAGGAGTCCCACTGTGGCAGTGCCTTCCCGCGGCTGG 3387  
Db 3301 GATGACCTGAACCTGCACAGGAAGGAGTCCCACTGTGGCAGTGCCTTCCCGCGGCTGG 3360  
QY 3388 GGGGACAGCACTGCAGGACACAGGAGGAGCGAGGGGTCACTGCTCAGAAATTCACAGCC 3447  
Db 3361 GGGGACAGCACTGCAGGACACAGGAGGAGCGAGGGGTCACTGCTCAGAAATTCACAGCC 3420  
QY 3448 TTGAGGCTCTACAGTGAACCTGAACACAGAGAGTGTCTGGGAGATTGGAAGTCTTCTAT 3507  
Db 3421 TTGAGGCTCTACAGTGAACCTGAACACAGAGAGTGTCTGGGAGATTGGAAGTCTTCTAT 3480  
QY 3508 AACGGGACCTGGGGCAGCGTGGGAGGAGGAACATACACAGCCATAGCAGGCAATTGTG 3567  
Db 3481 AACGGGACCTGGGGCAGCGTGGGAGGAGGAACATACACAGCCATAGCAGGCAATTGTG 3540  
QY 3568 TGCAGGAGCTGGGCTGTGGGAGAAATGAGTGTGAGCTCGCCCTTTATCTAAGACA 3627  
Db 3541 TGCAGGAGCTGGGCTGTGGGAGAAATGAGTGTGAGCTCGCCCTTTATCTAAGACA 3600  
QY 3628 GGCTCTGTTTCATGTGGGTTGATGACATTCAGTGTCTTAAAGCATATCTCCATATGG 3687  
Db 3601 GGCTCTGTTTCATGTGGTGTGATGACATTCAGTGTCTTAAAGCATATCTCCATATGG 3660  
QY 3688 CAGTGCCTGTGTCCTCCATGGGACGGAAGATCTCCAGCCCGCAGAGACCTGTGATC 3747  
Db 3661 CAGTGCCTGTGTCCTCCATGGGACGGAAGATCTCCAGCCCGCAGAGACCTGTGATC 3720  
QY 3748 ACATGTGAAGATGAATAAGAGTCCGTGGAGGAGACACCGAGTCTCTGGAGAGTGGAG 3807  
Db 3721 ACATGTGAAGATGAATAAGAGTCCGTGGAGGAGACACCGAGTCTCTGGAGAGTGGAG 3780  
QY 3808 ATCTGGCAGCAGGCTCTCTGGGGCACAGTGTGTGATGACTCTGGGACCTGGCCGAGGCG 3867  
Db 3781 ATCTGGCAGCAGGCTCTCTGGGGCACAGTGTGTGATGACTCTGGGACCTGGCCGAGGCG 3840

QY 3868 GAAGTGGTGTGTACAGAGCTGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGAGCGCTTCG 3927  
Db 3841 GAAGTGGTGTGTACAGAGCTGGCTGTGGCTCTGCTCTGGCTGCCCTGAGGAGCGCTTCG 3900  
QY 3928 TTTGGCCAGGAACTGGAACCATCTGTTGGATGACATCGGTGCAAGGAAATGAGTCA 3987  
Db 3901 TTTGGCCAGGAACTGGAACCATCTGTTGGATGACATCGGTGCAAGGAAATGAGTCA 3960  
QY 3988 TTTCTATGGAGTGTCAACCCCAACCCCTGGGACACAGTGAAGTGTGACACAAAGGAGAT 4047  
Db 3961 TTTCTATGGAGTGTCAACCCCAACCCCTGGGACACAGTGAAGTGTGACACAAAGGAGAT 4020  
QY 4048 GTGGCGTGTGAGTGTCTGTGACAGTGTGCAATCACTGAATGCTTCAAGTGTGAGTCA 4107  
Db 4021 GCTGGCGTGTGAGTGTCTGTGACAGTGTGCAATCACTGAATGCTTCAAGTGTGAGTCA 4080  
QY 4108 GCATTTATTTATCCAGTATCTTTGGGCTCTCTGCTGCTTCTGTTTATTTCTTCTC 4167  
Db 4081 GCATTTATTTATCCAGTATCTTTGGGCTCTCTGCTGCTTCTGTTTATTTCTTCTC 4140  
QY 4168 ACGTGTGTGCGAGTTCAGAAACAAACATCTGCCCTCTCAGAGTTCACACCAAGGAGG 4227  
Db 4141 ACGTGTGTGCGAGTTCAGAAACAAACATCTGCCCTCTCAGAGTTCACACCAAGGAGG 4200  
QY 4228 GGTCTCTGTGAGAGATTTTATTCATGAGTGTGAGACCTGCTCAGAGAGGAGCCCA 4287  
Db 4201 GGTCTCTGTGAGAGATTTTATTCATGAGTGTGAGACCTGCTCAGAGAGGAGCCCA 4260  
QY 4288 CATGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4347  
Db 4261 CATGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACA 4320  
QY 4348 TCCTGTGTGGAGTCTTCTGCTCTCTGAGGCCACAAA 4386  
Db 4321 TCCTGTGTGGAGTCTTCTGCTCTCTGAGGCCACAAA 4359  
RESULT 3  
AAS01217  
ID AAS01217 standard; cDNA; 3670 BP.  
XX  
AC AAS01217;  
XX  
AC AAS01217;  
DT 04-JUL-2001 (first entry)  
XX  
DE DNA encoding human secreted protein, POLY8.  
XX  
KW Human secreted protein; therapeutic; diagnostic; human; cancer; ss.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 113..3475  
FT CDS /\*tag= a  
FT /product= "Human secreted protein POLY8"  
XX  
PN WO200119856-A2.  
XX  
XX 22-MAR-2001.  
XX  
XX 13-SEP-2000; 2000WO-US25106.  
XX  
PR 13-SEP-1999; 99US-0153629.  
PR 16-SEP-1999; 99US-0154520.  
PR 20-SEP-1999; 99US-0154762.  
PR 13-OCT-1999; 99US-0159231.  
PR 12-SEP-2000; 2000US-0659634.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
XX Shimkets RA, Fernandes E, Herrmann JL, Liu X, Yang M, Boldog FL;  
PI  
XX WPI; 2001-244781/25.  
DR

DR P-PSDB; AAU00396.  
XX New POLYX polypeptide useful for treating or preventing a POLYX  
PT associated disorder, e.g. cancer  
XX  
XX  
PS Claim 9; Page 25-29; 152pp; English.  
XX  
XX The sequence represents the coding sequence of human secreted protein,  
CC POLYX. POLYX nucleic acids, polypeptides and antibodies to POLYX can be  
CC used for treating or preventing a POLYX associated disorder in a subject,  
CC preferably a human. These can be used in the manufacture of a medicament  
CC for treating a syndrome associated with a human disease selected from a  
CC POLYX-associated disorder, where the therapeutic is a POLYX polypeptide,  
CC a POLYX nucleotide or a POLYX antibody. They may also be used to screen  
CC for a modulator of activity, or latency, or predisposition to a POLYX  
CC associated disorder, e.g. cancer.  
XX  
SQ Sequence 3670 BP; 947 A; 752 C; 1073 G; 898 T; 0 other;  
  
Query Match 74.8%; Score 3459.8; DB 22; Length 3670;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3464; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
  
Qy 1114 GATGAGCAGATTGGAACTCGGACTAGCAGATGGAAGTACAAATTGTTACGGGAGAGTA 1173  
Db 200 GATGAGCAGATTGGAACTCGGACTAGCAGATGGAAGTACAAATTGTTACGGGAGAGTA 259  
  
Qy 1174 GAGGTGAGAAATCATGAACACTGGTGGACAAATATGTGACCAACTGGAAGAAATCAACAA 1233  
Db 260 GAGGTGAGAAATCATGAACACTGGTGGACAAATATGTGACCAACTGGAAGAAATCAACAA 319  
  
Qy 1234 GCCCTTGTGGTGTGTAAGCAGCTAGGATGTCGGTTTCACGGCTTTTGGCAGCTCGTCTGCT 1293  
Db 320 GCCCTTGTGGTGTGTAAGCAGCTAGGATGTCGGTTTCACGGCTTTTGGCAGCTCGTCTGCT 379  
  
Qy 1294 AAACCTAGTAATGAAGCTAGAGACATTTGGATAACACCATATCTTGGCACTGGGAATGAG 1353  
Db 380 AAACCTAGTAATGAAGCTAGAGACATTTGGATAACACCATATCTTGGCACTGGGAATGAG 439  
  
Qy 1354 TCAGCTCTCTGGGACTGCACATATGATGAAAGCAAGCAACATGCTTCCGGAAGATCA 1413  
Db 440 TCAGCTCTCTGGGACTGCACATATGATGAAAGCAAGCAACATGCTTCCGGAAGATCA 499  
  
Qy 1414 GATGCTGAGTAATTTGTTGATAAGGCAGATCTGGACCTTAAGCTTTGTCGGGCTCAT 1473  
Db 500 GATGCTGAGTAATTTGTTGATAAGGCAGATCTGGACCTTAAGCTTTGTCGGGCTCAT 559  
  
Qy 1474 AGCCCTCTTATGGAGATTGGAGTGAATACCAAGAGAGTGGGGACATCTGTGTCTAT 1533  
Db 560 AGCCCTCTTATGGAGATTGGAGTGAATACCAAGAGAGTGGGGACATCTGTGTCTAT 619  
  
Qy 1534 GACAGATGGACACAAAGGAATGCAGCTGTGTGTGTAACAAATTCGGGATGTGGAAGCCT 1593  
Db 620 GACAGATGGACACAAAGGAATGCAGCTGTGTGTGTAACAAATTCGGGATGTGGAAGCCT 679  
  
Qy 1594 ATGCATGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTTGGCTGGATGAC 1653  
Db 680 ATGCATGTGTTTGGTATGACCTATTTTAAAGAACATCAGGACCTATTTTGGCTGGATGAC 739  
  
Qy 1654 GTTCTTCGATTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAG 1713  
Db 740 GTTCTTCGATTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAG 799  
  
Qy 1714 CATAAATCTGTACACAGAGAGATGTGATTTGTAACCTGCTCAGGTGATGCAACATGGGGC 1773  
Db 800 CATAAATCTGTACACAGAGAGATGTGATTTGTAACCTGCTCAGGTGATGCAACATGGGGC 859  
  
Qy 1774 CTGAGCTGGTGGGGCGGAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGA 1833  
Db 860 CTGAGCTGGTGGGGCGGAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGA 919  
  
Qy 1834 CGGTGGGCACAGTGTGTGATGACGGCTGGAACTGAAGCTGACGCTGTGGTGTGTAGC 1893  
Db 1834 CGGTGGGCACAGTGTGTGATGACGGCTGGAACTGAAGCTGACGCTGTGGTGTGTAGC 1943

Db 920 CGGTGGGCACAGTGTGTGATGACGGCTGGAACGGTAAAGCTGACGCTGTGTGTAGC 979  
Qy 1894 CAGCTGGGACTGCCATCTTCTATCATTTGSCATGGGCTTGGGAAACGCTTCTACAGATAT 1953  
Db 980 CAGCTGGGACTGCCATCTTCTATCATTTGSCATGGGCTTGGGAAACGCTTCTACAGATAT 1039  
Qy 1954 GGAATAATTTGGCTCGATGATGTTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCATGC 2013  
Db 1040 GGAATAATTTGGCTCGATGATGTTTCTCTGTGATGAGATGAGTCAAGTCTCTGGTCATGC 1099  
Qy 2014 AGGAACAGTGGTGGGGAATAATCACTGCAGTCAAGTGAAGATGTTGGAGTGTCTGT 2073  
Db 1100 AGGAACAGTGGTGGGGAATAATCACTGCAGTCAAGTGAAGATGTTGGAGTGTCTGT 1159  
Qy 2074 TCTGATGATCGGATATGAGAGTGTGGGCTTGGGAGCAGCAGGCTGTCTGGAATA 2133  
Db 1160 TCTGATGATCGGATATGAGAGTGTGGGCTTGGGAGCAGCAGGCTGTCTGGAATA 1219  
Qy 2134 GTTGAAGTGAATGTCACGGGTGCCGTGGGAATTTCTGTGTCTTAATGCTGGGGAATGAAC 2193  
Db 1220 GTTGAAGTGAATGTCACGGGTGCCGTGGGAATTTCTGTGTCTTAATGCTGGGGAATGAAC 1279  
Qy 2194 ATTGCTGAAGTGTGTCAGGCAACTTGAATGTGGGCTGCAATCAGGCTCTCCAGAGAG 2253  
Db 1280 ATTGCTGAAGTGTGTCAGGCAACTTGAATGTGGGCTGCAATCAGGCTCTCCAGAGAG 1339  
Qy 2254 CCTCATTTACAGAAAGAACATTTACACATCTTTAATGTCGAAATTCCTGGCTGCACTGGAGG 2313  
Db 1340 CCTCATTTACAGAAAGAACATTTACACATCTTTAATGTCGAAATTCCTGGCTGCACTGGAGG 1399  
Qy 2314 GAAGCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACACTCGCTGTCTTAAAT 2373  
Db 1400 GAAGCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACAGACACTCGCTGTCTTAAAT 1459  
Qy 2374 ATGGAAGCAAGTTTCTGCTCAGCCACAGCCAGCCAGCCAGCTGGTGGAGCTGATATG 2433  
Db 1460 ATGGAAGCAAGTTTCTGCTCAGCCACAGCCAGCCAGCCAGCTGGTGGAGCTGATATG 1519  
Qy 2434 CCCTCTCTGAGAGTGTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2493  
Db 1520 CCCTCTCTGAGAGTGTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1579  
Qy 2494 GATTTCTCTCTTTCATGCTGCGCAATGTGCTGCGAGAAATTTAAATGTTGGAGATGCCATA 2553  
Db 1580 GATTTCTCTCTTTCATGCTGCGCAATGTGCTGCGAGAAATTTAAACTGTGGAGATGCCATA 1639  
Qy 2554 TCTCTTTCTGTTGGAGATCCTTTTGGAAAGGAATGGTCTAACTTGGCCCGGAAAGTTTC 2613  
Db 1640 TCTCTTTCTGTTGGAGATCCTTTTGGAAAGGAATGGTCTAACTTGGCCCGGAAAGTTTC 1699  
Qy 2614 CAGTGTGAAGGAGTGAACACTCACCCTTTCATTTATGCCCATTTGTTCAACATCCGGAAGAC 2673  
Db 1700 CAGTGTGAAGGAGTGAACACTCACCCTTTCATTTATGCCCATTTGTTCAACATCCGGAAGAC 1759  
Qy 2674 ACTTGTATCCACAGCAGAGAAATTTGGAGTGTCTGTTCGCCGATATACAGATGTCCGACTT 2733  
Db 1760 ACTTGTATCCACAGCAGAGAAATTTGGAGTGTCTGTTCGCCGATATACAGATGTCCGACTT 1819  
Qy 2734 GTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAACGCTGTGGACACTGGGGC 2793  
Db 1820 GTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAACGCTGTGGACACTGGGGC 1879  
Qy 2794 TCACATGTGTGACACCCACTGGGACCCAGAAAGATGCCGTGTTCTATGACAGACAGCTCAGC 2853  
Db 1880 TCACATGTGTGACACCCACTGGGACCCAGAAAGATGCCGTGTTCTATGACAGACAGCTCAGC 1939  
Qy 2854 TGTGGACTGTCTCTCAACCCAGAGGAAATATATTGGAGAAAGAGTGTCTGTGTG 2913  
Db 1940 TGTGGACTGTCTCTCAACCCAGAGGAAATATATTGGAGAAAGAGTGTCTGTGTG 1999  
Qy 2914 TGGGACACAGGTTTCAATTTAGGGAATGAGTCACTTCTCGATTAATGTCTCAATGACA 2973  
Db 2000 TGGGACACAGGTTTCAATTTAGGGAATGAGTCACTTCTCGATTAATGTCTCAATGACA 2059

QY 2974 GTTCTTGGAGCACTCCCTGTATCCATGGAATACCTCTCTGTGATCTGCACAGGAAGC 3033  
DB 2060 GTTCTTGGAGCACTCCCTGTATCCATGGAATACCTCTCTGTGATCTGCACAGGAAGC 2119  
QY 3034 CTGACCCAGCACTGTTTCCATGCTCGCAATATGATCTGACCCATATTTGTCTGCAGTT 3093  
DB 2120 CTGACCCAGCACTGTTTCCATGCTCGCAATATGATCTGACCCATATTTGTCTGCAGTT 2179  
QY 3094 CCAGAGGGCAGTCTTGTATCTCTTAGAGGACAAAGCGCTCGCCCTAGTGGATGGGAC 3153  
DB 2180 CCAGAGGGCAGTCTTGTATCTCTTAGAGGACAAAGCGCTCGCCCTAGTGGATGGGAC 2239  
QY 3154 AGCCGCTGTCCCGGAGAGTAGATCTATCACAGCGGCTTCTGGGCAACCATCTGTGAT 3213  
DB 2240 AGCCGCTGTCCCGGAGAGTAGATCTATCACAGCGGCTTCTGGGCAACCATCTGTGAT 2299  
QY 3214 GACGGCTGGACCTGACGATGCCAGTGGTGTGTCAAAAGCTGGGCTGTGAGTGGCC 3273  
DB 2300 GACGGCTGGACCTGACGATGCCAGTGGTGTGTCAAAAGCTGGGCTGTGAGTGGCC 2359  
QY 3274 TTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGGTACAGGGCCCATCTGGCTGGATGAC 3333  
DB 2360 TTCAATGCCACGGTCTCTGCTCACTTTGGGAGGGGTACAGGGCCCATCTGGCTGGATGAC 2419  
QY 3334 CTGAACCTGCACAGGAGTCCCACTTGTGGCAGTGGCCCTTCCCGGGCTGGGGCAG 3393  
DB 2420 CTGAACCTGCACAGGAGTCCCACTTGTGGCAGTGGCCCTTCCCGGGCTGGGGCAG 2479  
QY 3394 CAGACTGCAGGACAGGAGGAGCGAGGGTATCTGCTCAGAAATTCACAGCCTTGAGG 3453  
DB 2480 CAGACTGCAGGACAGGAGGAGCGAGGGTATCTGCTCAGAAATTCACAGCCTTGAGG 2539  
QY 3454 CTCTACAGTGAACCTGAACACAGAGAGTGTGCTGGAGATTGGAAGTCTTCTATAACGGG 3513  
DB 2540 CTCTACAGTGAACCTGAACACAGGAGTGTGCTGGAGATTGGAAGTCTTCTATAACGGG 2599  
QY 3514 ACCTGGGGCAGCGTGGCAGAGGAGCAATCACCACAGCCATAGCAGGCATTTGTGTCAGG 3573  
DB 2600 ACCTGGGGCAGCGTGGCAGAGGAGCAATCACCACAGCCATAGCAGGCATTTGTGTCAGG 2659  
QY 3574 CAGCTGGGCTGTGGGAGAGTGGAGTTGTGAGCTCCGCCCTTATCTAAGACAGGCTCT 3633  
DB 2660 CAGCTGGGCTGTGGGAGAGTGGAGTTGTGAGCTCCGCCCTTATCTAAGACAGGCTCT 2719  
QY 3634 GGTTCATGTGGTGGATGATCACTAGTGTCTTAAACGCATATCTCCATATGSCAGTGC 3693  
DB 2720 GGTTCATGTGGTGGATGATCACTAGTGTCTTAAACGCATATCTCCATATGSCAGTGC 2779  
QY 3694 CTGCTGCCCATGGGAGCGAAGATCTCCAGCCACAGACAGACCTGGATCACATGT 3753  
DB 2780 CTGCTGCCCATGGGAGCGAAGATCTCCAGCCACAGACAGACCTGGATCACATGT 2839  
QY 3754 GAACATAGATAAGAGTGGTGGAGGAGACACCGAGTGTCTGGGAGAGTGGATCTGG 3813  
DB 2840 GAACATAGATAAGAGTGGTGGAGGAGACACCGAGTGTCTGGGAGAGTGGATCTGG 2899  
QY 3814 CACGACGGCTCCTGGGACACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGCGGGAAGTG 3873  
DB 2900 CACGACGGCTCCTGGGACACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGCGGGAAGTG 2959  
QY 3874 GTGTGTACAGAGTGGGCTGTGGCTGTGCTGTGCTGCTGCTGAGGGAGCGTTCGTTTGGC 3933  
DB 2960 GTGTGTACAGAGTGGGCTGTGGCTGTGCTGTGCTGCTGCTGAGGGAGCGTTCGTTTGGC 3019  
QY 3934 CAGGGAACCTGGAACCATCTGTTGGATGACATGGGTGCAAGGAAATGAGTCAATTTCTA 3993  
DB 3020 CAGGGAACCTGGAACCATCTGTTGGATGACATGGGTGCAAGGAAATGAGTCAATTTCTA 3079  
QY 3994 TGGGACTGTACGCCAAACCTCTGGGAGAGAGTGTGCTGTGACACAAAGAGATGCTGGC 4053  
DB 3080 TGGGACTGTACGCCAAACCTCTGGGAGAGAGTGTGCTGTGACACAAAGAGATGCTGGC 3139

QY 4054 GTGAGTGTCTCGACAGTCGCTGAAATCACTGAATGCCTCTCAGCTCATTTAGCACTT 4113  
DB 3140 GTGAGTGTCTCGACAGTCGCTGAAATCACTGAATGCCTCTCAGCTCATTTAGCACTT 3199  
QY 4114 ATTTTATCCAGTATCTTTGGGCTCTCTCTGTTCTGTTTATTTCTTATTTCTCACCTGG 4173  
DB 3200 ATTTTATCCAGTATCTTTGGGCTCTCTCTGTTCTGTTTATTTCTTATTTCTCACCTGG 3259  
QY 4174 TGCCGAGTTTCAGAAAACAAAACATCTGCCCTCAGAGTTTCAACCAAGAGAGGGTCT 4233  
DB 3260 TGCCGAGTTTCAGAAAACAAAACATCTGCCCTCAGAGTTTCAACCAAGAGAGGGTCT 3319  
QY 4234 CTGAGAGAGAAATTTATTCATGATGGAGACCTGCTCAAGAGAGAGGCCCATGTTGG 4293  
DB 3320 CTGAGAGAGAAATTTATTCATGATGGAGACCTGCTCAAGAGAGAGGCCCATGTTGG 3379  
QY 4294 ACAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACATCGCTG 4353  
DB 3380 ACAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGATGCTAGCGACACATCGCTG 3439  
QY 4354 TTGGGAGTTCTTCTGCTCTGAAGCCACAAAATGACTTTTAGACTTCCAGGGCTCACCG 4413  
DB 3440 TTGGGAGTTCTTCTGCTCTGAAGCCACAAAATGACTTTTAGACTTCCAGGGCTCACCG 3499  
QY 4414 ATCAACCTCTAAATATCTTTGAAGGAGACACAACTTTTAAATCAATAAGAGAACTCA 4473  
DB 3500 CTCAACCTCTAAATATCTTTGAAGGAGACACAACTTTTAAATCAATAAGAGAACTCA 3559  
QY 4474 AGTTGCCCTATGAAAACCTTGTCCAAATAACATTTCTTGAACATAGGAGAACAGCTAAA 4533  
DB 3560 AGTTGCCCTATGAAAACCTTGTCCAAATAACATTTCTTGAACATAGGAGAACAGCTAAA 3619  
QY 4534 TTGATAAAGACTGGTGATAATAAAATTTGAATTATGTATATCACTGTATAA 4584  
DB 3620 TTGATAAAGACTGGTGATAATAAAATTTGAATTATGTATATCCCTGTATAA 3670

RESULT 4  
AAS44609  
ID AAS44609 standard; DNA; 2800 BP.  
XX  
AC AAS44609;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human full-length polynucleotide sequence #34.  
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;  
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;  
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;  
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
KW cytoskeletal; antirheumatic; antithrombotic; vasotropic; antiparkinsonian;  
KW antibacterial; immunosuppressive; vasodilator; antidiabetic; antiallergic;  
KW neuroprotective; osteopathic; analgesic; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200164834-A2.  
XX  
XX 07-SEP-2001.  
XX  
XX 26-FEB-2001; 2001WO-US04926.  
XX  
XX 28-FEB-2000; 2000US-0515126.  
XX 18-MAY-2000; 2000US-0577409.  
XX 17-JUN-2000; 2000US-0597707.  
XX 14-JUL-2000; 2000US-0616807.  
XX 19-SEP-2000; 2000US-0664541.  
PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
PI Drmanac R;  
XX WPI; 2001-589862/66.  
DR P-PSDB; AAU27709.  
XX Novel polypeptides and nucleic acids obtained from cDNA libraries  
PT prepared from various human tissues for diagnosis, treatment of  
PT cancer, neurological, inflammatory disorders and for use in arrays for  
XX detection  
PS Claim 1; SEQ ID No 34; 153pp; English.  
XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and  
CC contig polynucleotides encoding polypeptides of the invention. The DNA  
CC and protein sequences are useful for the treatment, diagnosis and  
CC prevention of various types of disorder in a mammalian subject such as a  
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers  
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such  
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
CC bowel disease. The sequences exhibit activity relating to angiogenesis,  
CC cell proliferation, cell differentiation, stem cell growth factor,  
CC activin or inhibin. Therefore, they can be used to manipulate stem cells  
CC in culture to give rise to neuroepithelial cells that can be used to  
CC augment or replace cells damaged by illness, accidental damage or genetic  
CC disorders. The sequences may also be used for regeneration of bone,  
CC cartilage, tendons and ligaments and in tissue repair and burn healing.  
CC Note: Some sequences for this patent did not form part of the printed  
CC specification, but were obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2800 BP; 697 A; 607 C; 813 G; 683 T; 0 other;  
Query Match 59.7%; Score 2761.2; DB:22; Length 2800;  
Best Local Similarity 99.7%; Pred. No.0;  
Matches 2766; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1821 GTACTTTCAAGGACGGTGGGCGACAGTGTGTGATCACGGCTGGAACAGTAAAGCTGCAGC 1880  
DB 27 GTGACTTTTCAGGACCGTGGGCGACAGCGTGTGATCACGGCTGGAACAGTAAAGCTGCAGC 86  
QY 1881 TGTGTGTGTAGCCAGCTGGAGTGGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGC 1940  
DB 87 TGTGTGTGTAGCCAGCTGGAGTGGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGC 146  
QY 1941 TTCTACAGATATGAAAAATTTGGCTCGATGATGTTTCTCTGTGATGGAGATGAGTCAGA 2000  
DB 147 TTCTACAGATATGAAAAATTTGGCTCGATGATGTTTCTCTGTGATGGAGATGAGTCAGA 206  
QY 2001 TCTCTGGTTCATCGAGAACAGTGGGTGGGGAATTAATGACTCAGTCAAGTGAAGATGT 2060  
DB 207 TCTCTGGTTCATCGAGAACAGTGGGTGGGGAATTAATGACTCAGTCAAGTGAAGATGT 266  
QY 2061 TGGAGTGTATCTTCTGTGATGATCGGATATGGAGCTGAGGCTGTGGGTGGGAAAGCAGCAG 2120  
DB 267 TGGAGTGTATCTTCTGTGATGATCGGATATGGAGCTGAGGCTGTGGGTGGGAAAGCAGCAG 326  
QY 2121 GTGTGCTGGAAAGTTGAGTGAATGTCAGGGTCCGTTGGGAATTTCTGTGTCTAATGG 2180  
DB 327 GTGTGCTGGAAAGTTGAGTGAATGTCAGGGTCCGTTGGGAATTTCTGTGTCTAATGG 386  
QY 2181 CTGGGGAATGAACATTGCTGAAGTTGTTTGGCAGCAACTTGAATGGGTCTGCAATCAG 2240  
DB 387 CTGGGGAATGAACATTGCTGAAGTTGTTTGGCAGCAACTTGAATGGGTCTGCAATCAG 446  
QY 2241 GGTCTCCAGAGAGCCTCAATTTACAGAAAGAACATTACACATCTTAATGTCGAATTTCTGG 2300

DB 447 GGTCTCCAGAGAGCGCTCATTTACAGAAAGAACATTACACATCTTAATGTCGAATTTCTGG 506  
QY 2301 CTGCACCTGGAGGGGAGCGCTCTCTCTGGATTGTATAGATGGGAGTGGAAACACACTGC 2360  
DB 507 CTGCACCTGGAGGGGAGCGCTCTCTCTGGATTGTATAGATGGGAGTGGAAACACACTGC 566  
QY 2361 GTGTCATTAAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCGACGCCAGCTGGT 2420  
DB 567 GTGTCATTAAATATGGAAGCAAGTTTGTATCTGCTCAGCCACAGCGACGCCAGCTGGT 626  
QY 2421 TGGAGCTGATATGCGCTCTCTGAGACGTGTGAACTGAAACATGACAGACACATGGCGCTC 2480  
DB 627 TGGAGCTGATATGCGCTCTCTGAGACGTGTGAACTGAAACATGACAGACACATGGCGCTC 686  
QY 2481 TGTCTGTCTATCTGATTCTCTCTTCATGCTGCCAATGTGCTGCAGAGAAATTAATTTG 2540  
DB 687 TGTCTGTCTATCTGATTCTCTCTTCATGCTGCCAATGTGCTGCAGAGAAATTAATCTG 746  
QY 2541 TGGAGATGCCATATCTCTTTCTGTGGGAGTCACTTTTGGAAAAGGGAATGGTCTTAAC 2600  
DB 747 TGGAGATGCCATATCTCTTTCTGTGGGAGTCACTTTTGGAAAAGGGAATGGTCTTAAC 806  
QY 2601 GCGCCAAAAGTTCCAGTGTGAAGGAGTGAACACTCACTTGCATTTATGCCCCATTTGTCA 2660  
DB 807 GCGCCAAAAGTTCCAGTGTGAAGGAGTGAACACTCACTTGCATTTATGCCCCATTTGTCA 866  
QY 2661 ACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTGTGAGTGTCTGTCTCCGATATAC 2720  
DB 867 ACATCCGGAAGACACTTGTATCCACAGCAGAGAAGTGTGAGTGTCTGTCTCCGATATAC 926  
QY 2721 AGATCTCCGACTTGTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 2780  
DB 927 AGATCTCCGACTTGTGAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAACGTGCT 986  
QY 2781 TGGACACTGGGCTCACTGTGTGACACCCACTGGGACCCAGAGAATGCCGTGTTCTATG 2840  
DB 987 TGGACACTGGGCTCACTGTGTGACACCCACTGGGACCCAGAGAATGCCGTGTTCTATG 1046  
QY 2841 CAGACAGCTCAGCTGTGGGACTGCTCTCAACACAGAGGAGAAATATATTGGAGAAG 2900  
DB 1047 CAGACAGCTCAGCTGTGGGACTGCTCTCAACACAGAGGAGAAATATATTGGAGAAG 1106  
QY 2901 AAGTGTCTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTCGATAA 2960  
DB 1107 AAGTGTCTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTCGATAA 1166  
QY 2961 CTGTCAATGACAGTTCCTTGGAGACCTCCCTGTATCCATGGAATACTGTCTCTGTAT 3020  
DB 1167 CTGTCAATGACAGTTCCTTGGAGACCTCCCTGTATCCATGGAATACTGTCTCTGTAT 1226  
QY 3021 CTGCACAGAACCTTGACCCAGCCACTGTTTCCATGCCCTCGCAATGTATCTGACCCATA 3080  
DB 1227 CTGCACAGAACCTTGACCCAGCCACTGTTTCCATGCCCTCGCAATGTATCTGACCCATA 1286  
QY 3081 TTTGTCTGCAGTTCAGAGGCGCAGTGTGATCTGCTTTAGAGACAAACGGCTCCGCT 3140  
DB 1287 TTTGTCTGCAGTTCAGAGGCGCAGTGTGATCTGCTTTAGAGACAAACGGCTCCGCT 1346  
QY 3141 AGTGAATGGGACACGCGCTGTCCGGGAGAGTAGAGATCTATCAGCAGCGGCTTCTGGG 3200  
DB 1347 AGTGAATGGGACACGCGCTGTCCGGGAGAGTAGAGATCTATCAGCAGCGGCTTCTGGG 1406  
QY 3201 CACCATCTGTGATGACGCTGGGACCTGAGCGATGCCACGTGTTGTGTCAAAAGCTGG 3260  
DB 1407 CACCATCTGTGATGACGCTGGGACCTGAGCGATGCCACGTGTTGTGTCAAAAGCTGG 1466  
QY 3261 CTGTGAGTGGCTTCAATGCCACGCTCTCTGCTCACTTTTGGGAGGGGTGAGGGCCCAT 3320  
DB 1467 CTGTGAGTGGCTTCAATGCCACGCTCTCTGCTCACTTTTGGGAGGGGTGAGGGCCCAT 1526  
QY 3321 CTGGCTGGATGACCTGAACCTGCACAGGAACGAGTCCCACTTGTGGCAGTGCCTTCCCG 3380

Db	1527	CTGGCTGGATGACCTGAACTGCACAGGAATGGAGTCCACACTTGTGGCGAGTGCCCTTCCCG	1588
Qy	3381	CGGCTGGGGGCAGCAGCTGCAGGCACAAAGGAGGAGCGAGGGTCACTGCTCGCAAAAT	3440
Db	1587	CGGCTGGGGGCAGCAGCTGCAGGCACAAAGGAGGAGCGAGGGTCACTGCTCGCAAAAT	1646
Qy	3441	CACAGCCTTTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT	3500
Db	1647	CACAGCCTTTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTGCTGGGAGATTGGAAGT	1706
Qy	3501	CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAGAACATCACACACGCCATACGAG	3560
Db	1707	CTTCTATAACGGGACCTGGGGCAGCGTCGGCAGGAGGAGAACATCACACACGCCATACGAG	1766
Qy	3561	CATTGTGTGCAGCAGCTGGCTGTGGGAGAAATGGAGTTGTCTCAGCCTGCCCTTTATC	3620
Db	1767	CATTGTGTGCAGCAGCTGGCTGTGGGAGAAATGGAGTTGTCTCAGCCTGCCCTTTATC	1826
Qy	3621	TAAGACAGGCTCTGGTTTCATCTGGGTGGATGACATTCAGTGTCTTAAACGCATATCTC	3680
Db	1827	TAAGACAGGCTCTGGTTTCATCTGGGTGGATGACATTCAGTGTCTTAAACGCATATCTC	1886
Qy	3681	CATATGTCAGTGCCTGTCTGCCCATGGGAGGAGAAATCTCCAAGCCACGACAGAGAC	3740
Db	1887	CATATGTCAGTGCCTGTCTGCCCATGGGAGGAGAAATCTCCAAGCCACGACAGAGAC	1946
Qy	3741	CTGGATCACATGTGAAGATAGAATAAGAGTGCCTGGAGGAGACACGAGTGCCTGGGAG	3800
Db	1947	CTGGATCACATGTGAAGATAGAATAAGAGTGCCTGGAGGAGACACGAGTGCCTGGGAG	2006
Qy	3801	AGTGGAGATCTGTCACGCGAGGCTCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGC	3860
Db	2007	AGTGGAGATCTGTCACGCGAGGCTCTGGGGCACAGTGTGTGATGACTCCTGGGACCTGGC	2066
Qy	3861	CGAGGCGGAAGTGGTGTCTCAGCAGCTGGCTGTGGCTCTCTCTGCTGCCCTGAGGGA	3920
Db	2067	CGAGGCGGAAGTGGTGTCTCAGCAGCTGGCTGTGGCTCTCTCTGCTGCCCTGAGGGA	2126
Qy	3921	CGCTTCTGTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGGGTGCCAAAGGAAA	3980
Db	2127	CGCTTCTGTTGGCCAGGGAACTGGAACCATCTGGTTGGATGACATGGGTGCCAAAGGAAA	2186
Qy	3981	TGAGTCAATTTCTATGGAATGTACGCCAAACCCCTGGGGACAGAGTGACTGTGGACACAA	4040
Db	2187	TGAGTCAATTTCTATGGAATGTACGCCAAACCCCTGGGGACAGAGTGACTGTGGACACAA	2246
Qy	4041	GGAAGATGCTGGCGTGAGTGCTCGGACAGTCGCTGAAATCACTGATGCCCTCCTCAGG	4100
Db	2247	GGAAGATGCTGGCGTGAGTGCTCGGACAGTCGCTGAAATCACTGATGCCCTCCTCAGG	2306
Qy	4101	TCATTTAGCACTTTATTTATCCAGTATCTTTGGGCTCCTCTCTCGTGGTCTGTGTTATCT	4160
Db	2307	TCATTTAGCACTTTATTTATCCAGTATCTTTGGGCTCCTCTCTCGTGGTCTGTGTTATCT	2366
Qy	4161	ATTTCTCAGCTGGTGGCGAGTTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACAG	4220
Db	2367	ATTTCTCAGCTGGTGGCGAGTTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACAG	2426
Qy	4221	AAGGAGGGGTTCTCTCGAGGAGAAATTTATTCATCAGATGGAGACCTGCCCTCAGAGAGA	4280
Db	2427	AAGGAGGGGTTCTCTCGAGGAGAAATTTATTCATCAGATGGAGACCTGCCCTCAGAGAGA	2486
Qy	4281	GGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAG	4340
Db	2487	GGACCCACATGGGACAAGAACCTCAGATGACACCCCCAACCATGGTTGTGAAGATGCTAG	2546
Qy	4341	CGACACATCGCTGTTGGGAGTTCTTCTCGCTCTCTGAAAGCCACAAAATGACTTTTACATTC	4400
Db	2547	CGACACATCGCTGTTGGGAGTTCTTCTCGCTCTCTGAAAGCCACAAAATGACTTTTACATTC	2606
Qy	4401	CAGGGCTCACGATCAACCTCTAAATATCTTTGAAGGAGACACAACTTTTAAATGAAT	4460
Db	2607	CAGGGCTCACGATCAACCTCTAAATATCTTTGAAGGAGACACAACTTTTAAATGAAT	2666

Qy	4461	AAGAGAAGTCAAGTTGCCCTATGGAACACTTGTCAAATAAACAATTTCTTGAACAATAG	4520
Dd	2667	AAGAGAAGTCAAGTTGCCCTATGGAACACTTGTCAAATAAACAATTTCTTGAACAATAG	2726
Qy	4521	GAGAACAGCTAAATTGATTAAGACCTGGTCATTAATAAAAAATTGAATTATGTATATCACTGT	4580
Dd	2727	GAGAACAGCTAAATTGATTAAGACCTGGTCATTAATAAAAAATTGAATTATGTATATCACTGT	2786
Qy	4581	TAAAAAAAAAAAAA 4594	
Dd	2787	TAAAAAAAAAAAAA 2800	
 RESULT 5 AAI58649			
ID	AAI58649	standard; cDNA; 3716 BP.	
XX	AC	AAI58649;	
XX	DT	22-OCT-2001 (first entry)	
XX	DE	Human polynucleotide SEQ ID NO 852.	
XX	KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;	
XX	KW	peripheral nervous system; neuropathy; central nervous system; CNS;	
XX	KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
XX	KW	anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
XX	KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
XX	KW	leukaemia; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200153312-A1.	
XX	PD	26-JUL-2001.	
XX	PF	26-DEC-2000; 2000WO-US34263.	
XX	PR	21-JAN-2000; 2000US-0488725.	
XX	PR	25-APR-2000; 2000US-0552317.	
XX	PR	09-JUL-2000; 2000US-0598042.	
XX	PR	19-JUL-2000; 2000US-0620312.	
XX	PR	03-AUG-2000; 2000US-0653450.	
XX	PR	14-SEP-2000; 2000US-0662191.	
XX	PR	19-OCT-2000; 2000US-0693036.	
XX	PR	29-NOV-2000; 2000US-0727344.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
XX	PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;	
XX	PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;	
DR	DR	WPI: 2001-442253/47.	
DR	DR	P-PSDB: AAM39493.	
XX	PT	Novel nucleic acids and polypeptides, useful for treating disorders	
XX	PT	such as central nervous system injuries -	
XX	PS	Claim 1: SEQ ID NO 852: 10078pp; English.	

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAK42213) with neurotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

XX Sequence 3716 BP; 993 A; 730 C; 1060 G; 933 T; 0 other;

Query Match 31.3%; Score 1448.2; DB 22; Length 3716;  
Best Local Similarity 65.4%; Pred. No. 0;  
Matches 2179; Conservative 0; Mismatches 1128; Indels 24; Gaps 3;

Qy	1002	GCTGTGATGGTCTCTCTGCTCCGGTAATGAATCTTTCTTTGGGAGCTCAGACATTCGGG	1061
Db	107	GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAGACATTTTGTCAACTTGAGTCCCTT	166
Qy	1062	AACCGTCAATTTGACTCTCTTCATCAAAACGATGCTGCTGTGATCTGCTCAGATGGAGC	1121
Db	167	CACCATTACTGTGGTCTTACTTCTCAGTCCCTGTTTGTCCACCAAGTTCTCTTGGAGGAA	226
Qy	1122	AGATTGTGAAGCTGGGACTAGCAGATGGAAGTAACAAATTTTCAGGGAGAGTAGAGTGGAG	1181
Db	227	AGACAGGAGCTGAGGCTAGTGGATGGTGAACAAAGTGTAGCGGAGAGTGGAGTGA	286
Qy	1182	AATTCATGAACAGTGGTGGCAATATGTACACAGAACTGGAAGAAATGAACAGCCCTTGT	1241
Db	287	AGTCCAGGAGGTGGGAAACGGTGTGAATAATGGCTGGAGCATGGAAGCGGTCTCTGT	346
Qy	1242	GGTTTGAAGCAGTATGATGCTCCGTTCCAGGCTCTTTGGCAGTCTGCTGTCTAAACCTAG	1301
Db	347	GATTTGTAAACAGCTGGGATGTCACAACTGCTATCAAGCCCTCGATGGGCTTAATTCAG	406
Qy	1302	TAATGAAGCTAGACATTTTGGATAAAGCAGATATTTTGCACCTGGGAATGAGTCACTCT	1361
Db	407	TGCAGTTCTGCAGCAATTTGGATGATCATGTTTCTTGTCTGGATGAGTCACTCT	466
Qy	1362	CTGGAGCTGCACATATGATGAAGAAAGCAAGC---GAACATGCTTTCGGAAGATCAGATGC	1418
Db	467	TTGGATTTGCAACATATGATGGGAAAGCATAGTAACTGCTACTACCAACAAAGATGC	526
Qy	1419	TGGAGTAATTTGTTCTGATAAGGAGATCTGACCTAAGGCTGTGCGGGCTCATAGCCC	1478
Db	527	TGGAGTGACCTGCTCAGATGATCAATTTTGGAAATGAGGCTGAGCGGTGGAGGAATAT	586
Qy	1479	CTGTTATGGGAGATGGAGTGCAATACCAAGGAGTGGGGAGTGTGTCTCATGACAG	1538
Db	587	GTGTTCTGGAAGATAGAGTCAAAATCCAAAGGAGGTGGGGAACAGTGTGTGATGATA	646
Qy	1539	ATGGAGCAAGGAATGAGCTGTGTGTGTAAACAAATGGGATGTGGAAGCCCTATGCA	1598
Db	647	CTTCAACATATAGTATGATCTGTCATTTGTAGAACAACTTGAATGTGGAAGTGTGTG	706
Qy	1599	TGTGTTTGGTATGACTATTTTAAAGAGCATCAGGACCTATTTGGCTGGATGAGTTTC	1658
Db	707	TTTCTCTGGTTCATTAATTTGGAGAGAGCTCTGGACCAATCTGGTTTGTATGATCTTAT	766
Qy	1659	TTGCAATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGAAAGCATAA	1718
Db	767	ATGCAACGGAATGAGTCACTCTGGAATCTGGAACATCAAGGATGGGGAAGCATAA	826
Qy	1719	TTGTGTACACAGAGGATGATTTGAACCTGCTCAGGATGATGCAACATGGGGCTGAG	1778
Db	827	CTGTGATCATCTGAGGATGTGGAGTGTATTTGTTCAAGGAGGACGATCTGAGCCTGAG	886
Qy	1779	GCTGTGGGGGACAGACCCCTGCTCGGGAAGCTGAGGCTGACTTTTCAAGGAGCGTG	1838
Db	887	ACTGTTAGATGAGTCACTGAATGTTTCAGGAAGATTTAGAGTGTAGATTTCCAAAGGAATG	946
Qy	1839	GGGCACAGTGTGTATGACGGCTGGAACAGTAAAGCTGCAGCTGTGTGTGTAGCCAGCT	1898
Db	947	GGGACAAATATGTGATGACGGCTGGACAGTACGATCTGCTGTGGCATCCAGCAACT	1006
Qy	1899	GGACTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGGAA	1958

Db	1007	GGGATGTCCAACCTCCGCTCAGACCCATTTGGTCTAGTTAACGCCAGTAAAGGATTTGGACA	1066
Qy	1959	AATTTGGCTCCGATGATGTTTCTCTGTGATGGAGATGAGTACAGATCTCTGTCTCATGAGGAA	2018
Db	1067	CATCTGGCTTGACACGGTTTCTTTCAGGAGCATGAACCTGCTGTCTGGCAATGAACA	1126
Qy	2019	CAGTGGTGGGAAATAATAGTGCAGTCACAGTGAAGATTTGGAGTATCTGTTCTCA	2078
Db	1127	CCATCAATGGGAAAGCATTTATGCAATCACAATGAAGATGCTGGCGTACATGTTCTGA	1186
Qy	2079	TGATCGGATATGAGAGCTGAGGCTTGTGGTGGAGAGCAGAGTGTGTCTGGAAGATTTGA	2138
Db	1187	TGGATCAGATCTGGAGCTTAAGACTTAGAGGTGGAGGAGCCCTGTCTGTGGACAGTTGA	1246
Qy	2139	SGTGAATCTCCAGGCTCCGCTGGGAATTTCTGTGTCTAATGCTGGGCTGGGAATGAACATTC	2198
Db	1247	SGTGGAGATTCAGAGACTGTGTAGGGAAGTGTGTGACAGAGCTGGGACTGAAAGAACG	1306
Qy	2199	TGAAGTTCTTTCAGGCAACTTGAATGTGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA	2258
Db	1307	TGATGTGCTTTCAGGCAACTTGAATGTGGTCTGCAATCAGGGTCTCCAGAGAGCCTCA	1366
Qy	2259	TTTCACAGAAAGAACATTTACACATCTTAATGTCGAATTTCTGTGCTGACATGGAGGGAAGC	2318
Db	1367	CTCCAAATCCAGGCAACAAACACATGGCTGTTCTAAGTAGCTGTAAACGAAATGAAC	1426
Qy	2319	CTCTCTCTGGGATTTATACGATGGGAGTGGGAAACAGACTGCTGCTCATTTAAATATGA	2378
Db	1427	TTCTTTTGGGACTTCAAGAACTTGGCAATGGGTGGACTTACCTGTGATCATTAAGA	1486
Qy	2379	AGCAAGTTTGTATCTCTCAGCCACAGGCAAGGCTGGTGGAGTGTATGATGCTGATGCTG	2438
Db	1487	AGCCAAATTAACCTCTCAGCCACAGGAAACCCAGACTGGTGGAGGAGCATTTCCCTG	1546
Qy	2439	CTCTGGAGCTGTTGAAGTGAACATGCAGACATGGCGCTCTGTCTGTGATTTCTGATTT	2498
Db	1547	TTCTGGAGCTGTTGAAGTGAAGCATGGTGCACAGTGGGCTCCATCTGTGATTTCCGACTT	1606
Qy	2499	CTCTCTTCACTGCTGCAATGTCTGTCAGAGAAATTAATTTGTCGAGATGTCATATCTCT	2558
Db	1607	CTCTCTGAGAGCTGCCAGCTTCTATGACAGGAATTAACAGTGTGACAGTGTCTCTAT	1666
Qy	2559	TTCTCTGGGAGATCACTTTTGGAAAGGAATGTTCTAACTTGGCGCCGAAAGTTCAGTG	2618
Db	1667	CCTGGGGGAGCTCACTTTGGAGAGGAAATGGACAGATCTGGCTGAAAGATTTCCAGTG	1726
Qy	2619	TGAAGGAGTGAACATCACTTGCATTTATGCCCAATTTTCAACATCCGGAAGACACTTG	2678
Db	1727	TGAGGAGATGAGTCCCATCTTTCACTCTGCCAGTAGCACCCGCCAGAGAACTTG	1786
Qy	2679	TATCCACAGCAGAGAAAGTTGGAGTTGCTGTTTCCGATATACAGATGTCGCACTTTGTGA	2738
Db	1787	TAGCCACAGCAGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGTGAA	1846
Qy	2739	TGGCAAAATCCCAGTGTGAGCGGCAAGTGGAGATCAACCTGCTTGGACACTGGGGCTCACT	2798
Db	1847	TGGCAAGACCCGCTGTAGGGCAGAGTGCAGCTCAAAACGCTTGTGCTGGGATCCCT	1906
Qy	2799	GTGTACACCCACTGGGACCCAGAGATGCCGCTTCTATGACACAGCTCAGCTGTGG	2858
Db	1907	CTGTAACTCTCACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTTAAATGTGG	1966
Qy	2859	GACTCTCTCTCAACACAGGAGAAATATATTTGGAGAAAGAGTGTTCGTGTGTGGG	2918
Db	1967	AGTTGCCCTTCTACCCAGGAGGAGCAGGTTTTCGAAAGGAATGTCAGATCTGGAG	2026
Qy	2919	ACACAGGTTTCTATGCTTATGGGAATGAGTCACTTCTTGGATAACTGTCAAAATGACAGTTCT	2978
Db	2027	GCATATGTTTCACTGCAGCTGGGAGTGGAGACATGGAGATTTGCTCTGTAACTGCTCT	2086
Qy	2979	TGGAGCACCTCCCTCTATCCATGGAATACTGCTCTGTGATCTGCACAGGAGCCTGAC	3038



Qy 1002 GCTTGGTGGTCTCTCCGCTAATGAATCTTTCTTTGGGACTGCAGACATTCGGG 1061  
Db 107 GCTACTTGAAGACTCTGGATCTGCTACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166  
Qy 1062 AACCGTCAATTTTGGACTCTCTTCATCAAAAGCATGTGTCTGTGATCTGCTCAGATGGAGC 1121  
Db 167 CACCACTACTGTGTCTTACTTCTCAGTGCCTGTTTGTCCACCACTTCTCTTTGGAGAAC 226  
Qy 1122 AGATTTGGAACTGGCATAGCAGATGAAGTAACAATTTTTCAGGGAGAGTAGAGGTGAG 1181  
Db 227 AGACAAGGAGCTGAGGTAGTGGATGTGGAACAAGGTGAGCGGAGAGTGGGAAGTGA 286  
Qy 1182 AATTCAATGAACAGTGGTGGCAATATGTGACCAGAATCTGGAAGAATGAACAAGCCCTGTG 1241  
Db 287 AGTCCAGGAGGTGGGAAGCGGTGTGTAATAGTGGTGGAGCATGGAAGCGGTCTCTGT 346  
Qy 1242 GGTTTGTAAGCAGTAGAGTGCCTGCTCAGCGTCTTTGGCAGTCTGCTGTGTAACCTAG 1301  
Db 347 GATTTGTAACCCAGCTGGGATGTCCAATCTGCTATCAAAAGCCCTGGATGGGCTAATTCAG 406  
Qy 1302 TAATGAAGCTAGACACATTTTGGATAAACAGCATATCTTGGCACTGGGAATGAGTCAGCTCT 1361  
Db 407 TGCAGGTTCTGGAGCGATTTGGATGGATCATGTTCTTGTCTGGGAATGAGTCAGCTCT 466  
Qy 1362 CTGGGACTGCACATATGATGAAAAGCAAAAGC --- GAACATGCTTCCGAAGATCAGATGTC 1418  
Db 467 TTGGATTTGCAACATGATGGATGGGGAAGCATAGTAACCTACTCACCAACAAGATGC 526  
Qy 1419 TGGAGTAATTTGCTGATAGGAGATCTGGACCTAAGGCTTTGTGGGGCTCTATAGCCC 1478  
Db 527 TGGAGTGACCTGCTCAGATGATCCAAATTTTGGAAATGAGGCTGACGGCTGGAGGGAATAT 586  
Qy 1479 CTGTTATGGGAGATGGAGGTGAATACCAAGGAGAGTGGGGGACTGTGTCATGACAG 1538  
Db 587 GTGTTCTGGGAAGATAGATCAAAATCCAAAGGACGGTGGGGAACAGTGTGTGATGATAA 646  
Qy 1539 ATGGAGCACAAAGGATGAGCTGTGTGTGTAACAATTTGGATGTGGAAGCCCTATGCA 1598  
Db 647 CTCAACATAGATCATGCTGTCTATTTGTAGACAATCTGAATGTGGAAGTGTGTCAG 706  
Qy 1599 TGTGTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTGGCTGGATGAGCTTTC 1658  
Db 707 TTTCTCTGGTTTCATCTAATTTTGGAGAGGCTCTGGACCAATCTGGTTGTGATCTTAT 766  
Qy 1659 TTGATTGGAAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAGCATAA 1718  
Db 767 ATGCAACGGAATGAGTCAAGTCTCTGGAATCTGCAACATCAAGATGGGGAAGCATAA 826  
Qy 1719 TTGTTACACAGAGGATGTGATTTGTAACCTGCTCAGGTGATGCAACATGGGCGCTGAG 1778  
Db 827 CTGTGATCATGCTGAGGATGCTGGAGTGTATTTGCTCAAGGGAGCAGATCTGAGCCTGAG 886  
Qy 1779 CTGTTGGGCGGCAACACCCCTGCTGGGAAGACTGAGGTGTACTTTTCAAGGACGGTG 1838  
Db 887 ACTGTTAGATGGAGTCACTGAATGTTCAAGGAAGATTAGAAGTGAAGATTCCAAGGAGAATG 946  
Qy 1839 GGGCACAGTGTGATGACGCTGGAACAGTAAAGCTCAGCTGTGGTGTGTAGCCAGCT 1898  
Db 947 GGGACAATATGTGATGACGCTGGGACAGTTTACGATGCTGCTGTGGCATGCAAGCAACT 1006  
Qy 1899 GGACTGCCCATCTTCTATCTATTGGATGGGTCTGGGAACCGTCTTACAGGATATGAAA 1958  
Db 1007 GGGATGTCCAACTCCGCTCAGACCATTTGGTTCGAGTTTAAAGCCAGTAAAGGATTTGGACA 1066  
Qy 1959 AATTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAAGATCTCTGCTCATGCAAGAA 2018  
Db 1067 CATCTGGCTTGACAGCGTTTCTTGCCAGGGACATGAACCTGTGTCTGCTGCAATGTAACA 1126  
Qy 2019 CAGTGGGTGGGGAATATGACTGCAGTCAAGTGAAGATGTTGGAGTGATCTGTCTGA 2078  
Db 1127 CCATGAATGGGGAAGCATTTATGCAATCACAATGAAGATGCTGGCGTGACATGTCTGA 1186

Qy 2079 TGCATCGGATATGGAGCTGAGGCTTGTGGTGAAGCAGCAGGTGCTGTGGAAGTGA 2138  
Db 1187 TGGATCAGATCTGGAGCTTAACACTTAGAGGTGGAGCAGCCGCTGTGCTGGACAGTTGA 1246  
Qy 2139 GGTGAATGTCAGGCTGCGGTGGGAATTTCTGTGTCTAATGGCTGGGAATGAACATGTC 2198  
Db 1247 GGTGAGATTCAGAGACTGTTTAGGAAGGTGTGTGACAGAGCTGGGACTGGAAGAAGC 1306  
Qy 2199 TGAAGTTGTTTTCAGGCAACTTGAATGTTGGTCTGCAATCAGGCTCCAGAGAGCCCTCA 2258  
Db 1307 TGATGTGTTTTTTCAGGCACTGGGATGTGGATCTGCATCAAAACATCTTATCAAGTGA 1366  
Qy 2259 TTTTCACAGAAAACATTTACACATCTTAAATGTCGAAATTTCTGGCTCAGCTGGAGGGAAGC 2318  
Db 1367 CTCAAAATCCAGGCAACAACACATGCTGTTTCTAAGTAGCTTAACGGGAATGAAC 1426  
Qy 2319 CTCCTCTCTGGATTTGATATACGATGGGAGTGGAAACAGACTGGTGTCTATTTAAATATGA 2378  
Db 1427 TTTCTTTTGGGACTGCAAGAACTGGCAATGGGCTGGACTTACCTGTGATCACTATGAAGA 1486  
Qy 2379 AGCAAGTTTGTATCTGCTCAGCCACAGCAGCCAGGCTGGTTGGAGCTGATATGCCCTG 2438  
Db 1487 AGCCAAAATTAACCTGCTCAGCCACAGGGAACCCAGACTGGTTGGAGGAGACATTCCTG 1546  
Qy 2439 CTCGGAGCTGTGAAGTGAACATGCAGACACATGGCGCTCTGTCTGTGATCTGTGATTT 2498  
Db 1547 TTTCTGGAGCTGTGAAGTGAAGCATGTTGACACGTGGGCTCCATCTGTGATTCGGACTT 1606  
Qy 2499 CTCCTCTCATGCTGCCAATGTGCTGCAGAGAAATTAATTTGTGGAGATGCCATATCT 2558  
Db 1607 CTCTCTGGAAGCTGCCAGCTTCTATGCAAGGAATTTACAGTGTGACAGTTGTCTCTAT 1666  
Qy 2559 TTTCTGTGGAGATCACTTTTGGAAAAGGAATGGTCTAACTTGGCGCGAAAAGCTTCCAGTG 2618  
Db 1667 CTGGGGGAGCTCACTTTGGAGGGAATGGACAGATCTGGGCTGAAGAATTCAGATG 1726  
Qy 2619 TGAAGGAGTGAACCTCACTTTGCAATATGCCCCATTTTCAACATCCGGAAGACACTTG 2678  
Db 1727 TGAGGAGATGAGTCCCATCTTTTCACTCTGCCAGTAGCACCCGCCCAAGAACTTG 1786  
Qy 2679 TATCCACAGCAGAGAAATGGAGTTGTGTTCCCGATATACAGATGTCGACTTGTGAA 2738  
Db 1787 TAGCCACAGCAGGAGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGA 1846  
Qy 2739 TGGCAATCCCACTGTGACGGCAAGTGGAGATCAACGTCTTGGACACTGGGGCTCACT 2798  
Db 1847 TGGCAAGACCCGCTGTGAGGCGAGATGAGCTCAAAAGCTTTGGTCTGGGATCCCT 1906  
Qy 2799 GTGTGACACCCACTGGGACCCAGAAAGATGCCGCTGTCTATGACAGACAGCTCAGCTGG 2858  
Db 1907 CTGTAACTCTCACTGGGACATAGAAAGTCCCATGTTCTTTGCCAGCAGCTTAAATGTGG 1966  
Qy 2859 CACTGCTCTCAACCCAGAGGAGAAATATATTGAGAAAGAAAGTGTGCTGTGGGG 2918  
Db 1967 AGTTCCCTTTCTACCCAGGAGGAGCAGGTTTTTGGAAAGGAATGTCAGATCTGGAG 2026  
Qy 2919 ACACAGGTTTCAATCTTAGGGAATGAGTCACTTCTGGAATACTGTCAAAATGACAGTTCT 2978  
Db 2027 GCATATGTTTCACTGCACTGGGACTGAGCAGCAGATGGAGATTTGCTGTAACTGCTCT 2086  
Qy 2979 TGGAGCAGCTCCCTGATTCATGGAATACTCTCTCTGTGATCTGCAGAGAGGCTGAC 3038  
Db 2087 AGGTGCTTCAATATCTCTTCAAGCAAGTGGCTCTGTAAATCTGCTCAGAGAAACAGTC 2146  
Qy 3039 CCAGCCACTGTTTCCATGCTCCGCAATGTATCTGACCCATATTTGTGTGAGTTCCAGA 3098  
Db 2147 CCAACACTGCTCTGTCGAATTTCACTGCTTTTGGGCCCAACAGGCGCTTACCATTCCAGA 2206  
Qy 3099 GGGCAGCTTGTGATCTGCTTAGAGGACAAACGGCTCGCCCTAGTGGATGGGAGCAGCGC 3158  
Db 2207 AGAAGTGTGTGGCTGCATAGAGAGTGGTCAACTTCCCTGGTGAATGGAGAGGTGCG 2266  
Qy 3159 CTGTCCCGGAGAGTAGAGATCTATCAGACGCGCTTCTGGGCGACCATCTGTGATGACGG 3218

Db 2267 CTGTGCTGGAGAGTAGATCTATCATGAGGCTCTCTGGGACCATCTGTGATGACAG 2326  
QY 3219 CTGGGACCTGAGCGATGCCACAGTGTGTCTAAAAGCTGGGCTGTGGAGTGGCTTCAA 3278  
Db 2327 CTGGGACCTGAGTGTGATGCCACAGTGTGTGCACACAGCTGGGCTGTGGAGAGCCATTAA 2386  
QY 3279 TGCCACGGTCTCTGCTCACTTTTGGGAGGGGTGAGGGCCCATCTGTGCTGGATGACCTGAA 3338  
Db 2387 TGCCACCTGTTCTGCTCACTTTTGGGAGGAACAGGGCCCATCTGTGCTGGATGAGATGAA 2446  
QY 3339 CTGCACAGAAAGGAGTCCCACTTTGGGAGTGGCAGTCCCTTCCCGGGCTGGGGCAGCACCA 3398  
Db 2447 ATGCAATGAAAAGAAATCCGCAATTTGGCAGTGGCAATTCACACGGCTGGGGCAGCAAAA 2506  
QY 3399 CTGCAGGCACAGGAGGAGCGAGGGTCACTCTGCTCAGAAATTCACAGCTTGGAGGCTCTA 3458  
Db 2507 TTGCAGGCACAGGAGATGCGGGAGTTATCTGCTCAGAAATTCATGCTCTGAGACTGAC 2566  
QY 3459 CAGTGAACCTGAAACACAGAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAAGGGGACCTG 3518  
Db 2567 CAGTGAAGCCAGCAGAGAGGCGCTGTGCAGGGCTCTGGAAGTTTTTTACAATGGAGCTTG 2626  
QY 3519 GGGCAGCTGGCAGAGAGAAATCACCACAGCCATAGCAGGATTTGTGTGAGGCGAGCT 3578  
Db 2627 GGGCAGCTGTGGCAAGTAGCATGTCTGAAACCACTGTGGGTGTGTGTGAGGCGAGCT 2686  
QY 3579 GGGCTGTGGGAGAGTGGTGTGAGCTCGCCCTTTATCTAAGCAGAGCTCTGGTTT 3638  
Db 2687 GGGCTGTGACAGAAAGGAAATCAACCCCTGCATCTTTAGACAGGCCATGTCCATTCC 2746  
QY 3639 CATGTGGGTGGATGACATTCAGTGTCTTAAAGCGCATATCTCCATATGGCAGTGCCTGTC 3698  
Db 2747 CATGTGGGTGGCAATGTTTCAGTGTCCAAAGACCTGACACGCTGTGGCAGTGCCTATC 2806  
QY 3699 TGCCCATGGGAGGAGAAATCTCCAGCCACAGCAGAGAGACCTGTGATCATGTCAAGA 3758  
Db 2807 ATCTCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGACCTGTGATCATGTGACAA 2866  
QY 3759 TAGAATAAGAGTGCCTGGGAGAGACCGAGTGTCTCTGGGAGAGTGGAGATCTGGCAGCG 3818  
Db 2867 CRAATGAAGCTCAGGAGAGCCACTTCTGTTCTGGACGTGTGGAGATCTGGCATGG 2926  
QY 3819 AGGCTCTGGGACAGTGTGTGATGACTCTCTGGACCTGGCGGAGCGGAGTGTGTG 3878  
Db 2927 AGGTTCTGGGACAGTGTGTGATGACTCTTGGGACTTGGAGATGCTCAGTGTGTG 2986  
QY 3879 TCAGCAGCTGGGCTGTGCTCTGCTGCTGCTGCCCTGAGGAGCTTCGTTTGGCCAGG 3938  
Db 2987 TCACCACTGGCTGTGGTCCAGCTTTGAAGCATTTCAAGAGCAGAGTTTGGTCAGGG 3046  
QY 3939 AACTGGAACCATCTGGTTGGATGACATGGGTGCAAGGAAATGAGTCAATTTCTATGGGA 3998  
Db 3047 GACTGGACCATATGCTCAATGAAGTGAAGTCAAGGAAATGAGTCTCTCTTGGGA 3106  
QY 3999 CTGTCAGCCCAACCTCGGGACAGAGTGTGAGTGTGACACAGAGAGATGCTGGCGTGAG 4058  
Db 3107 TTGCTCTGCGACAGCTGGGGCCATAGTGTGAGTGTGGCAGCAGAGAGAGCTGCACTGAA 3166  
QY 4059 GTGCTCTGGACAGCTGCTGAA-----ATCACTGAATGCCCTCCTCAGG 4100  
Db 3167 TTGCACAGATATTTAGTGCAGAAACCCACAAAAGCCACACAGGTGGCTCATCCCG 3226  
QY 4101 TCATTTAGCACTTTATTTATCCAGTATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCT 4160  
Db 3227 TCAGTCATCTTTATTTGAGTGGGATCTCTGGGTTGTTCTGTGGCCATTTTCGTGCG 3286  
QY 4161 ATTTCTCAGTGTGCGGAGTTCAGAAACAAACATCTGCCCTCAGATTTTCACAG 4220  
Db 3287 ATTTATTTCTTTGACTTAAAGCGAGACAGAGACA---CGGGCTTGGAGTTTCTTCAAG 3343  
QY 4221 AAGGAGGGGTCTCTCGAGAGAGATTTATTTCCATGAGATGGAGACCTGCTCAAGAGAGA 4280

Db 3344 AGGAGAGAACTTAGTCCACCAAAATCAATACCGGAGATGAATTTGCTGCTGAATCAGA 3403  
QY 4281 GGACCCACATGGGACAAAGAACCTCAGATGAC 4311  
Db 3404 TGATCTGGACCTAATGAATTCCTCAGGAGGC 3434  
RESULT 7  
AAI60435/c  
ID AAI60435 standard; cDNA; 3834 BP.  
XX AAI60435;  
AC AAI60435;  
XX  
DT 22-OCT-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 4424.  
DE  
DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI: 2001-442253/47.  
DR P-PSDB; AAM41279.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
PS Claim 1; SEQ ID NO 4424; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;

Query Match	31.38;	Score 1448.2;	DB 22;	Length 3834;
Best Local Similarity	65.4%;	Pred. No. 0;		
Matches 2179;	Conservative	0;	Mismatches 1128;	Indels 24;
Gaps	3;			
Qy 1002	GCTTGATGGTCTCCCTGCTCCGGTAATGAATCTTTCTTTGGGACTGCACACATTCGGG	1061		
Db				
Db 3728	GCTACTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT	3669		
Qy 1062	AACCGTCAATTTTGACTGTCTTTCATCAAAAGATGTGTCTGTGATCTGCTCAGATGAGC	1121		
Db				
Db 3668	CACCATTACTGTGCTTACTTCTTCAGTGCCTGTTTGTCAACAGTTCTCTTGGAGAAC	3609		
Qy 1122	AGATTGGAATCGGACTAGCAGATGGAAGTAACAATTTGTTCAGSGAGAGTAGAGGTGAG	1181		
Db				
Db 3608	AGACAAGAGCTGAGGCTAGTGGATGGTGAANAACAGTGTACCGGGAGAGTGGAGGTAA	3549		
Qy 1182	AATTCAATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGAATGAACAAGCCCTTGT	1241		
Db				
Db 3548	AGTCCAGAGGAGTGGGAACGGTGTGTAATAATGGCTGGAGCATGGAAGCGGTCTCTGT	3489		
Qy 1242	GGTTTGAAGCAGTAGGATGTCGGTTCAGCGTCTTTGGCAGTCTGTCGTCTAAACCTAG	1301		
Db				
Db 3488	GATTTGAACAGCTGGGATGTCCAACCTGCTATCAAAAGCCCTGGATGGGTAAATCCAG	3429		
Qy 1302	TAATGAAGCTAGACACATTTGGATAAACAAGACATATCTTGCACTGGGAATGAGTCAGCTCT	1361		
Db				
Db 3428	TGCAGGTTCTGGAGCATTTGATGGATCATGTTCTTGCTGGGAATGAGTCAGCTCT	3369		
Qy 1362	CTGGGACTGCACATATGATGGAAGAAAGCAAGC --- GAACATGCTTCCGAAGATCAGATGC	1418		
Db				
Db 3368	TTGGGATTTGCAACATGATGGATGGGAAGACATAGTAACCTGTACTACCAACAAGATGC	3309		
Qy 1419	TGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGCGTTGTCGGGGCTCATAGCCC	1478		
Db				
Db 3308	TGGAGTGACCTGCTCAGATGATCCAATTTGGAATGAGGCTGACCGTGGAGGAATAT	3249		
Qy 1479	CTGTTATGGGAGATTTGAGGTGAATCCAAAGGAGAGTGGGGGACTGTGTCTCATGACAG	1538		
Db				
Db 3248	GTGTTCTTGGAGAATAGAGATCAAAATTTCCAAGGACGGTGGGGAACAGTGTGTGATGATA	3189		
Qy 1539	ATGGAGCACAAGGAATGCAGCTGTGTGTGTGAACAATTTGGGATGTGGAAAGCCTATGCA	1598		
Db				
Db 3188	CTTCAACATAGATCATGCATCTGTCATTTGTAGACAACTTGAATGTGGAAGTGTGTGAC	3129		
Qy 1599	TGCTTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTGGCTGGATGACGTTTC	1658		
Db				
Db 3128	TTTCTCTGGTTCACTCAATTTTGGAGAAGCTCTGGACCAATCTGGTTGATGATCTTAT	3069		
Qy 1659	TTGCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGGGGAAGCATAA	1718		
Db				
Db 3068	ATGCAACGGAAATGAGTCAAGCTCTCTGGAACCTGCAAAACATCAAGGATGGGGAAGCATAA	3009		
Qy 1719	TTGTGTACACAGAGAGTGTATTGTAACTGCTCAGGTGTCAACATCGGGGCTGAG	1778		
Db				
Db 3008	CTGTGATCATCTGAGGATCTGGAGTGATTTGCTCAAGGGAGCAGATCTGACGCTGAG	2949		
Qy 1779	GCTGGTGGGCGGACAAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG	1838		
Db				
Db 2948	ACTGGTAGATGGAGTCACTGAAATGTTCAAGGAAGATTAGAAGTGAAGTATTTCAAGGAATG	2889		
Qy 1839	GGGCACAGTGTGTATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTCTAGCCAGCT	1898		
Db				
Db 2888	GGGACAATATGTATGACGGCTGGACAGTTACGATGCTGCTGTGGCATGCAAGCAACT	2829		
Qy 1899	GGACTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAACGCTTCTACAGGATATGGAAA	1958		
Db				
Db 2828	GGGATGTCCAACCTGCCGTCACAGCAATTTGGTTCAGGTAAACGCCAGTAAGGATTTGGACA	2769		
Qy 1959	AATTTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCAGATCTCTCGGTCTATGCAGGAA	2018		
Db				
Db 2768	CATCTGGCTTGACAGCGTTTCTTGCCAGGGGACATGAACCTGCTGTCTGGCAATGTAAACA	2709		



CC Note: The sequence data for this patent did not form part of the printed  
specification.

XX  
SQ Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other:  
Query Match 31.3%; Score 1448.2; DB 22; Length 3834;  
Best Local Similarity 65.4%; Pred. No. 0;  
Matches 2179; Conservative 0; Mismatches 1128; Indels 24; Gaps 3;

```
QY 1002 GCTTCATGGTCTCTCTCGGTAATGAATCTTTCTTTGGGACTCAGACATTCGGG 1061
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3728 GCTACTTGAAGACTCTGGATCTGCTGACTTTCAGAGACATTTTGTCAACCTGAGTCCCTT 3669
QY 1062 AACCTCAATTTTGACTGTCTTCATCAAAACGATGTCTGTGATCTGCTCAGATGGAGC 1121
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3668 CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTTTGTCCACCACTTCTTTGGAGAAC 3609
QY 1122 AGATTGGAACGTGCACATAGCAGATGGAAGTAACAAATTTTTCAGGAGAGTAGAGTGAG 1181
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3608 AGACAAGGAGCTGAGGCTAGTGATGGTGAACAAAGTGTAGCGGAGAGTGAAGTGAA 3549
QY 1182 AATTCATGAACAGTGTGGACAATATGTACCAGACATGGAAGATGAACAGCCCTTGT 1241
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3548 AGTCCAGGAGGAGTGGGAACGGTGTGAATTAATGGCTGGAGCATGGGAAGCGGTCTCTGT 3489
QY 1242 GCTTTGTAAGCACTAGGATGCCCTTTCAGCGCTCTTTGGCAGTCTGCTGTCTAAACCTAG 1301
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3488 GATTTGTAAACAGCTGGGATGTCCTCAACTGCTATCAAGCCCTGGATGGCTTAATCCAG 3429
QY 1302 TAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTTCACATGGGAATGAGTCAGCTCT 1361
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3428 TGCAGGTTCTGACGCATTTGGATGATCATGTTCTTGTCTGGGAATGAGTCAGCTCT 3369
QY 1362 CTGGACATGCACATATGATGGAAGCAAGC ---GAACATGCTTCCGNAAGATCAGATGC 1418
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3368 TTGGGATTGCAACATATGATGGAAGCAAGCATAGTAACTACTACTCAACAAAGATGC 3309
QY 1419 TGGAGTAATTTGCTGTAAGGCAGATCTGGACCTAAGGCTGTGCGGGCTCATAGGCC 1478
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3308 TGGAGTACCTCTCAGATGGATCCATTTGGAAATGAGCTGACGCGTGGAGGAATAT 3249
QY 1479 CTGTTATGGAGATTTGGAGTGAATACCAAGGAGAGTGGGGATGTGTGTATGACAG 1538
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3248 GTGTTCTGGAAGATAGATCAAAATTCGAAGGCGGTGGGAACAGTGTGTATGATGATAA 3189
QY 1539 ATGGAGCAAGAGATGAGCTGTGTGTGAACAATTTGGATGTGGAACCCATATGCA 1598
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3188 CTTCACATAGATCATGATCTGTCTATTGTAGACAACCTTGAATGTGGAAGTGTGTGACAG 3129
QY 1599 TGTGTTTGGTATGACCTATTTTAAAGACATCAGGACCTATTTGGCTGGATGAGCTTTC 1658
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3128 TTTCTCTGTTTCATCTAATTTTGGAGAGGCTCTGGACCAATCTGTTTGTATGATCTTAT 3069
QY 1659 TTGCATTTGGAATAGTCAAAATATCTGGGATGTGAACACAGTGGATGGGGAAGACATAA 1718
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3068 ATGCAACGGAATAGTCAAGCTCTCTGGAATCTGAACATCAAGATGGAAGGAAGACATAA 3009
QY 1719 TTGTGTACACAGAGGATGTGATTTACCTGCTCAGGTGATGCAACATGGGGCTGAG 1778
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3008 CTGTGATCATGTGAGGATGTGGAGTGATTTGCTCAAGGGAGGAGATCTGAGCCTGAG 2949
QY 1779 GCTGTGGGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGACGGTG 1838
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2948 ACTGTAGATGAGTCACTGAAATGTTTCAGGAAGATTAAGAGTGAATTCAGAGGAGATG 2889
QY 1839 GGGCAGTGTGTATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCCAGCT 1898
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2888 GGGCAATATATGTATGACGGCTGGACAGTTACGATGTGTGTGTGGCATGCAAGCAACT 2829
QY 1899 GGACTGCCCATCTTCTATCATTTGGCATGGTCTGGAAACGCTTCTACAGGATATGGAAA 1958
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2828 GGGATGTCCAACTGCCGTACAGCCATTTGGTCAGTTTAAACGCCAGTAAGGGATTTGGACA 2769
```

```
QY 1959 AATTGGCTCGATGATGTTTCTCTGTGATGGAGATGAGTCAGATCTCTGTCATGCAGGAA 2018
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2768 CATCTGGCTTGACACCGTTTCTTGGCAGGACATGAACCTGCTGCTGCAATGTAACA 2709
QY 2019 CAGTGGGTGGGAAATAATAGCTTGCAGTCACAGTGAAGATGTTGGAGTATGTTCTGTA 2078
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2708 CCATGAATGGGAAAGCATTTATTGCAATCACAATGAAGATGCTGGCGTGACATGTTCTGA 2649
QY 2079 TGCATCGATATGGAGCTGAGGCTTGTGGTGAAGCAGCAGGTGCTGTGGAAAAATTGA 2138
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2648 TGGATCAGATCTGGAGCTTAAGACTTAGAGGTGGAGCAGCCGCTGTCTGCGACAGTTGA 2589
QY 2139 GGTGAATGCCAGGTCGCGTGGGAATTTCTGTGCTTAATGGCTGGGAAATGAACATGTC 2198
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2588 GGTGAGATTCAGACACTGTTAGGGAAGGTGTGCACAGAGCTGGGACTGAAAGAAGC 2529
QY 2199 TGAAGTTGTTGCAGGCAACTTGAATGTGGTCTCAATCAGGCTCTCCAGAGACCCCTCA 2258
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2528 TGATGTGTTTGCAGGCACTGGGATGTGGATCTGCACCTCAAAACATCTTATCAAGTGTGA 2469
QY 2259 TTTTCACAGAAACATTTACACATCTTAATGTCGAATTTCTGCTGCACCTGGAAGGGAAGC 2318
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2468 CTCGAAATCCAGGCAACAAACATGCTGTTCTAAGTAGCTGTACCGGAATGAAC 2409
QY 2319 CTCTCTCTGGGATTTATACGATGGGAGTGGAAACAGACATGCTGCTCATTTAAATATGA 2378
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2408 TTTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACCTGTGATCACTATGAAGA 2349
QY 2379 AGCAAGTTTGATCTCTCAGCCCAAGGAGCCAGGCTGGTTGGAGCTGATATGCCCTG 2438
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2348 AGCCAAAAATTTACTCTCTCAGCCCAAGGGAACCCAGACTGGTTGGAGGGGACATCCCTG 2289
QY 2439 CTCTGAGCTGTGAAGTGAACATGCACACATGCGGCTCTGCTGTGATTTCTGATTT 2498
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2288 TTTCTGAGCTGTGAAGTGAAGCATGGTGACAGTGGGGCTCCATCTGTGATTCGGACTT 2229
QY 2499 CTCTCTTCTGATGCTCAATGTCTGTGCAGAGAAATTAATTTGGAGATGCCATATCTCT 2558
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2228 CTCTCTGGAAGCTGCCAGCTTCTATGCAGGGAATTTACAGCTGTCACAGTTGTCTCTAT 2169
QY 2559 TTTCTGTGGAGATCACTTTGGAAAAAGGATGTTAACTTGGGCGGGAAGATTTCCAGTG 2618
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2168 CTTGGGGGAGCTCACTTTGGAGAGGAAATGGACAGATCTGGGCTGAAGAAATTTCCAGTG 2109
QY 2619 TGAAGGGATGAACCTCACTTTGCATTTATGCCCATTTGTTCAACATCCGGAAGACACTTG 2678
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2108 TGAGGACATGAGTCCCATCTTTCACTCTGCCAGTAGGACCCGCCCAAGAGAACTTG 2049
QY 2679 TATCCACAGCAGAGAGTTGGAGTTGTCTGTTCCCGATATACAGATGTCCGACTTGTGAA 2738
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2048 TAGCCACAGCAGGATGTTGGAGTAGTCTGCTCAAGATACACAGAAATTCGCTTGGTGA 1989
QY 2739 TGGAAATCCCAGTGTGAGGGCAAGTGAGATCAACGTGTTGGACACTGGGGCTCACT 2798
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1988 TGGCAAGACCCGCTGTGAGGGCAGAGTGGAGCTCAAAAGCTTTGGTCTGGCGATCCCT 1929
QY 2799 GTGTGACACCCACTGGGCCCCAGAGATGCCGCTGTTCTATGCAGACAGCTCAGCTGTGG 2858
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1928 CTGTAACCTCACTGGGACATAGAAGTGGCCATTTCTTTTGGCAGAGCTTAAATGTGG 1869
QY 2859 GACTGCTCTCTCAACACAGAGGAAATATATTGGAGAAAGATGTTGCTGTGTGGG 2918
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1868 AGTTGCCCTTTCTACCCAGGAGGAGCAGGTTTGGAAAAAGAAATGGTCAGATCTGGAG 1809
QY 2919 ACACAGGTTTCAATGCTTAGGGAATGAGTCACTTCTTGGAAATGTCAAAATGACAGTTCT 2978
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1808 GCATATGTTTCACTGCACTGGGACTGAGCAGCACATGGGAGATTTGCTGTAACTGCTCT 1749
QY 2979 TGGACACCTCCCTGATCCATGGAATACTCTCTGTGATGTCACAGAGGAGCTGAC 3038
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1748 AGGTGCTTCAATGATGCTTTCAGAGCAAGTGGCCCTCTGTAATCTGCTCAGGAAACAGTC 1689
QY 3039 CCAGGCCACTGTTTCCATGCTCGCAAAATGATATCTGACCCATATTTTGTCTGACAGTTCCAGA 3098
```

Db 1688 CCAAACTCTCTCGTCAATTCATCTCTTTGGGCCCCAACAGGCTACCATTCAG 1629  
QY 3099 GGGAGTGTCTTATCTCTAGGACAAACGGCTCCGCTAGTGTAGTGGGACAGCG 3158  
Db 1628 AGAAGTGTCTGGCTTGCATAGAGAGTGTCACTTCGCTGTAAATGAGAGGTG 1569  
QY 3159 CTGTGGCGGAGAGTAGAGATCTATCAGCAGCGCTCTTGGGGCACCATCTGTGATGACGG 3218  
Db 1568 CTGTGGCGGAGTAGAGATCTATCATGAGGCTCTTGGGGCACCATCTGTGATGACAG 1509  
QY 3219 CTGGAGCTGAGGATGCGCCACGTGTGTCTCAAAAGCTGGGTGTGGAGTGGCTTCAA 3278  
Db 1508 CTGGAGCTGAGTGTGCTTCAATTTGGGGAAGAACAGGCGCTCTGGCTGTGATGATGAA 1449  
QY 3279 TGCACAGCTCTGTCTCCTCCTTTGGGAGGCTCAGGCGCCATCTGGCTGTGATGACCTGAA 3338  
Db 1448 TGCACAGCTGTCTCTCCTTCAATTTGGGGAAGAACAGGCGCTCTGGCTGTGATGATGAA 1389  
QY 3339 CTGCACAGGAGGAGTCCCACTTTGTGGCAGTGTCCCTTCCCGCGCTGGGGGAGCAGCA 3398  
Db 1388 ATGCAATGGAAGAAATCCCGCATTTGGCAGTGCCTATTCACGCTGGGGGAGCAAAA 1329  
QY 3399 CTGCAGGACAGGAGGAGGAGGCTCTCTCTCAGAAATTCACAGCCTTGGGCTCTA 3458  
Db 1328 TTGCAGGACAGGAGGAGTGGGAGTATCTCTCAGAAATTCATCTCTGAGACTGAC 1269  
QY 3459 CAGTGAACCTGAAACAGAGAGTGTGTCTGGGAGTGTGAAGTCTCTATACCGGAGCTG 3518  
Db 1268 CAGTGAAGCCAGCAGAGGCTGTGGAGGCTGTGGAAGTCTTTTACAAATGGAGCTG 1209  
QY 3519 GGGCAGCTGGCAGGAGGAGGAGTCAACACAGGCTATAGCAGGATGTGTGAGGAGCT 3578  
Db 1208 GGGCAGCTGTGGCAAGAGTGTGAGTGTGAAACAGGCTGTGGGTGTGGTGTGAGGAGCT 1149  
QY 3579 GGGCTGTGGGAGATGGAGTGTGAGCTGTGAGCTGTGAGCTGTATCTAAGCAGGCTGTGTTT 3638  
Db 1148 GGGCTGTGACAGCAAGGAGGAGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 1089  
QY 3639 CATGTGGTGTGATGACATCTCAGTGTCTTAAACGATATCTTCCATATGGCAGTGTGCTGTC 3698  
Db 1088 CATGTGGTGTGACATCTCAGTGTCTTAAACGATATCTTCCATATGGCAGTGTGCTGTC 1029  
QY 3699 TGCCCCATGGGAGGAGGAGTCTCCAGCCAGCAGAGAGACCTGGATCATGTGAGA 3758  
Db 1028 ATCTCCATGGGAGGAGGAGTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 969  
QY 3759 TAGAATAGAGTGTGTGGAGGAGGAGTGTGAGCTGTGAGGAGTGTGAGCTGTGAGCT 3818  
Db 968 CAGATAGAGCTTCAGGAGGAGGAGGAGTGTGAGCTGTGAGGAGTGTGAGCTGTGAGCT 909  
QY 3819 AGGCTCTGGGAGGAGGAGTGTGAGCTGTGAGCTGTGAGGAGGAGGAGGAGGAGTGTG 3878  
Db 908 AGGTTCTGGGAGGAGGAGTGTGAGCTGTGAGCTGTGAGGAGGAGGAGGAGGAGTGTG 849  
QY 3879 TCAGCAGTGGGCTGTGCTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3938  
Db 848 TCACAACTGTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 789  
QY 3939 AACTTGAACCTGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3998  
Db 788 GACTTGGACCGATATGGCTCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 729  
QY 3999 CTGTACGCCCAACCTCTGGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4058  
Db 728 TTGTCTTCCAGGAGCTGTGGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 669  
QY 4059 GTGCTGTGAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 4100  
Db 668 TTGACAGATATTTTCACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609  
QY 4101 TCATTTAGCACTTTATTCAGTATCTTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4160

Db 608 TCAGTCACTCTTATTTGCACTGGGATCTTGGGGTGTGCTTTCGCGCATTTTCCTCGC 549  
QY 4161 ATTCTCACGTGTGCGGAGTTTCAGAAACAAACATCTGCCCTCAGAGTTTCAACACG 4220  
Db 548 ATTATCTTCTTCTGACTTAAAGGAGGAGACAGACA--GCGCTTTCAGTTTCTCTCAAG 492  
QY 4221 AAGGAGGGTCTCTCGAGGAGATTTATTCATGATGAGAGGAGGAGGAGGAGGAGGAG 4280  
Db 491 AGGAGAGAACTTGTAGTCCAGCAATTCATACCGGAGATGAATCTTGTGCTGAATGAGA 432  
QY 4281 GGACCCACATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4311  
Db 431 TGATCTGGACCTTAATGAATTCCTCAGAGGC 401  
RESULT 9  
ID AAS94922 standard; DNA; 3707 BP.  
XX AAS94922;  
XX AC  
XX DT  
XX DT  
XX 14-FEB-2002 (first entry)  
XX Human DNA sequence #177 expressed during foam cell differentiation.  
XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
XX cardiovascular disorder; coronary artery disease; gene therapy; ds.  
XX Homo sapiens.  
XX WO200177389-A2.  
XX 18-OCT-2001.  
XX PD  
XX XX  
XX PF 04-APR-2001; 2001WO-US11128.  
XX XX  
XX PR 05-APR-2000; 2000US-195106P.  
XX XX  
XX PA (INCY-) INCYTE GENOMICS INC.  
XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;  
XX Tai J;  
XX WPI; 2002-010925/01.  
XX Composition useful for diagnosis of conditions, disorders or diseases  
XX associated with atherosclerosis, comprises several polynucleotides that  
XX are differentially expressed in foam cell development  
XX Claim 1; Page 232-233; 315pp; English.  
XX The present invention relates to the isolation of human polynucleotide  
XX sequences that are differentially expressed during foam cell  
XX differentiation. The polynucleotide sequences of the invention or a  
XX composition comprising these polynucleotides are useful as a high  
XX throughput method for detecting altered expression of one or more  
XX polynucleotides in a sample. The polynucleotides can be used in the  
XX diagnosis of disorders associated with foam cell development such as  
XX atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
XX coronary artery disease. The polynucleotide sequences can also be used  
XX as PCR primers and probes. The polynucleotides of the invention are also  
XX useful in gene therapy. AAS94746-AAS95021 represent the human  
XX polynucleotide sequences of the invention which are differentially  
XX expressed during foam cell differentiation.  
XX Sequence 3707 BP; 984 A; 731 C; 1062 G; 930 T; 0 other;  
Query Match 30.5%; Score 1409.4; DB 24; Length 3707;  
Best Local Similarity 65.2%; Pred. No. 0;  
Matches 2176; Conservative 0; Mismatches 1131; Indels 28; Gaps 6;  
QY 1002 GCTGTGATGTCTCTCTGCTCGGTAATGAATCTTTTGGGAGTGTGAGGAGTGTGAGG 1061

Db 107 GCTACTGAAGACTCTGGATCTCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCCTT 166  
QY 1062 AACCGTCAATTTTGACTGTCTTCATCAAAACGATGTGTCTGTCTGCTCAGATGAGC 1121  
Db 167 CACCATTTACTGTGTCTTACTTCTCAGTCCCTGTTTGTCAACAGTCTCTCTTGAGGAAC 226  
QY 1122 AGATTTGGAACTCGCACTAGCAGATGAAGTAAACAATTTG-TCAGGAGAGTAGAGGTGA 1180  
Db 227 AGACAAGGAGCTGAGGCTAGTGGATGGTGAAGAAACAAGTGTAGCGGGAGAGTGAAGTGA 286  
QY 1181 GAATTCATGAACAGTGGTGGACAATATGTGACCAGAACTGGAAGATGAACAAGCCCTTG 1240  
Db 287 AAGTCCAGGAGGAGTGGGAAACGCTGTGTAAATAGCTGGAGCATGGAAGGCTCTCTG 346  
QY 1241 TGGTTTGAAGCAGCTAGGATGTCGTTACAGCGTCTTTGGCAGTCGTCGTAAACCTA 1300  
Db 347 TGATTTGTAACCAAGCTGGGATGTCCTAACTGCTATCAAAAGCCCTGGATGGGCTAAATCCA 406  
QY 1301 GTAATGAAGCTAGAGACATTTGGATAAACAAGCATATCTTGCACCTGGGAATGAGTCAAGCTC 1360  
Db 407 GTCAGGTTCTGGACGATTTGGATGGATGATCATGTTCTTCTGCTGGGAATGAGTCAAGCTC 466  
QY 1361 TCTGGGACTGCATATGATGGAAGAAAGCAAGC---GAACATGCTTCCGAAGATCAGATG 1417  
Db 467 TTTGGGATTGCAAAACATGATGAGTGGGAAAGCATAGTAACCTACTACCAACAAGATG 526  
QY 1418 CTGAGTAATTTGTTCTGATAAGCAGATCTGGACCTTAAGCCTTGTGGGGCTCATAGCC 1477  
Db 527 CTGAGTGACCTGCTCAGATGGATCCANTTTGGAAATGAGGCTGACCGCTGGAGGAAATA 586  
QY 1478 CTTGTATGGAGATTGGAGGTGAATACCAAGGAGAGTGGGGACTGTGTGTCATGACA 1537  
Db 587 TGTGTTCTGGAAGAATAGACATCAAAATTCGAAGCAGCTGGGGAACAGTGTGTGATGATA 646  
QY 1538 GATGGACCAAGGATGCAAGCTGTGTGTGTAACAAATTTGGGATGTTGGAAGCCTATGC 1597  
Db 647 ACTTCAACATAGATCATGATCATGTCTATTTGTAGACAACCTTGAATGGAAGTGTCTCA 706  
QY 1598 ATGTGTTGCTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGCTGGATGAGCCTTT 1657  
Db 707 GTTCTCTGTTCTAATTTTGGAGAGGCTCTGGACCAATCTGTTTGTGATGATCTTA 766  
QY 1658 CTTGATTTGGAATAGCTCAAAATATCTGGACATGTGGAACAGTGTGAACAGTGGATGGGAAAGCATA 1717  
Db 767 TATGCAACGGAATAGTCAAGCTCTCTGGAATGCAAAACATCAAGGATGGGAAAGCATA 826  
QY 1718 ATTGTGTACACAGAGGATGTGATTTGTAACCTGCTCAGTGATGCAACATGGGGCTGA 1777  
Db 827 ACTGTGATCATGTGAGGATGCTGGAGTGTATTTGCTCAAAAGGAGCAGATCTCAGGCTGA 886  
QY 1778 GCGTGTGGGGCGGAGCAACCGCTGCTCGGAAGACTGGAGGTGTACTTTCAAGGACGGT 1837  
Db 887 GACTGGTAGATGAGTCACTGAATTTTCAAGAGATTTAGAAGTGAATTTCAAGGAGAAAT 946  
QY 1838 GGGGCACAGTGTGTGATGACGGCTGGAAACAGTAAAGCTGACGCTGTGGTGTGAGCCAGC 1897  
Db 947 GGGGGACAATATGTGATGACGGCTGGGACAGTTACGATGCTGTGGCATGCAAGCAAC 1006  
QY 1898 TGGACTCCCATCTTCTATCATTTGGCATGGTCTGGGAACCGTCTTACAGGATATGGA 1957  
Db 1007 TGGGATGTCACCACTGCGCTCAGACCCATTTGTCGAGTTAAAGCCAGTAAGGGATTTGGAC 1066  
QY 1958 AAATTTGGCTCGATGATGTTTCTGTCATGGAGATGAGTCAGATCTCTGTCATGAGGA 2017  
Db 1067 ACATCTGGCTTGACAGGCTTTTGGCCAGGACATGAACCTGCTGCTGSCAATGTAAC 1126  
QY 2018 ACAGTGGGTGGGAAATAATGACTGACGTACAGTGAAGATGTTGGAGTCACTGTGTTCTG 2077  
Db 1127 ACCATGAATGGGAAGACATATTGTCATACATGAAGATGCTGCGCTGACATGTTCTG 1186  
QY 2078 ATGCATCGGATATGGAGCTGAGGCTTGTGGTGGAAACAGCAGGATGCTGGAAGATTG 2137  
Db 1187 ATGGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGGACCGGCTGTGCTGGGACAGTTG 1246

QY 2138 AGTGAATGTCCAGGGTCCGCTGGGAATTTCTGTGTCTAATGCTGGGAATCAACATTTG 2197  
Db 1247 AGTGGAGATTACAGACTGTTAGGGNAGTGTGTGACAGAGCTGGGACTGNAAGAAG 1306  
QY 2198 CTGAAGTTGTTTCAGAGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCAGAGAGCCTC 2257  
Db 1307 CTGATGTGGTTTCAGGACAGCTGGATGTGGATCTGCACTCAAAACATCTTATCAAGTGT 1366  
QY 2258 ATTTACAGAAAGACATTAACATCTTAATGTGCAATTTGCTGCTGCACTGGAGGGAAG 2317  
Db 1367 ACTCCAAAATTCAGGCAACAACACATGCGTGTCTTAAAGTAGCTGTAAACGGAATGAAA 1426  
QY 2318 CCTCTCTCTGGATTTGTATACGATGGAGTGGAAACAGACTGCGTGTCAATTTAAATATGG 2377  
Db 1427 CTCTCTTTGGGACTGCAAGAACTGGCAATGGGGTGGACTTACTGTGATCACTATCAAG 1486  
QY 2378 AAGCAAGTTTGAATCTGCTCAG---CCCACAGGACGCCAGGCTGTTGGAGCTCATATGCC 2435  
Db 1487 AAGCCAAAATTAACCTGCTCAGGCCACAGGGAAACCCAGACTGTTGGAGGGACATTC 1546  
QY 2436 CTGCTCTGGAGCTGTTGAAGTGAACATGACAGACATGCGGCTCTCTCTGTGATTTCTGA 2495  
Db 1547 CTGTTCTGAGAGCTGTTGAAGTGAAGCATGTGTGACACGCTGGGGCTCCATCTGTGATTCAGA 1606  
QY 2496 TTTCTCTCTCATCTGCTGCCAATGTGCTGTGACAGAAATTAATTTGTGGAGATGCCATATC 2555  
Db 1607 CTCTCTCTGGAAGCTGCCAGCGTCTTATGCGAGGAATTAACAGTGTGGCAGCTGTCTC 1666  
QY 2556 TCTTTCTGCTGGGAGATCATTGGAAGGAAATGGTCTAACTTTGGGCCGAAAGATTCCA 2615  
Db 1667 TATCCTGGGGAGCTCATTGAGAGGAAATGGACAGATCTGGCTGAAGAATTTCCA 1726  
QY 2616 GTGTGAGGAGTGAACACTCACCTTGCAATATGCCCATTTTCAACATCCGGAAGACAC 2675  
Db 1727 GTGTGAGGAGATGAGTCCCCTCTTCTCCTGCTGCCCCAGTACACCCCGCCAGGAAGAAC 1786  
QY 2676 TTGTATCCACAGCAGAGAAGTTGGAGTTGCTTCTCCCGATATACAGATGTCCTGACTTGT 2735  
Db 1787 TTGTAGCCACAGCAGGAGTGTGGAGTAGTCTGCTCAAGATACACAGAAATTCCTTGTGT 1846  
QY 2736 GAATGGCAATCCAGTGTGACGGGCAAGTGGAGATCAAGTGTGTTGGACACCTGGGGCTC 2795  
Db 1847 GAATGGCAAGCCCGTGTGAGGCGAGATGGAGCTCAAAACGCTTGGTGCCTGGGGATC 1906  
QY 2796 ACTGTGTGACACCCACTGGGACCCAGAGATGCCCGTTCTATGACAGACAGCTCACGCTG 2855  
Db 1907 CCTGTGTAACCTCTCAGTGGGACATAGAGATGCCCCATGTTCTTTGCCAGCAGCTTAATG 1966  
QY 2856 TGGGACTGCTCTCTCAACCCAGGAGGAAATATATTGGAGAAAGAAAGTGTCTGTGTG 2915  
Db 1967 TGGAGTTGCCCTTTCTACCCAGGAGGAGCAGCTTTTGGAAAAGAAATGGTCAGATCTG 2026  
QY 2916 GGGACAGAGTTTCAATGCTTAGGAAATGAGTCACTTCTGTGATCACTGTCAATGACAGT 2975  
Db 2027 GAGGCATATGTTTCACTGSCACTGGGACTGAGCAGACATGGGAGATTTCTCTGTAAATGC 2086  
QY 2976 TCTTGGAGCAGCTCCCTGTATCCATGCAATACTGCTCTGTGATCTGACAGCAAGCCT 3035  
Db 2087 TCTAGTGTCTTCAATGATGCTTCCAGAGCAAGTGGCCCTCTGTATCTGCTCAGGAACCA 2146  
QY 3036 GACCCAGCCACTGTTCCATGCTTCGCAAAATGTATCTGACCCATPATTTGTCTGCAAGTTC 3095  
Db 2147 GTCACAAACACTGTCTCTGTCGTAATTCATGCTCTTTGGGCCCAACAAGGCTACCATTC 2206  
QY 3096 AGAGGCGAGTGTGTGATCTGCTTAGAGGACAAACGCGCTCGCCCTAGTGTGATGGGACAG 3155  
Db 2207 AGAAGAAGTGTCTGCGCTGCAATAGAGTGTGTCAACTTTCGCTGTTAAATGAGGAGG 2266  
QY 3156 CCGCTGTGCGGGAGTAGAGATCTATCAGGAGCGCTTCTGGGGACCACTCTCTGTATGA 3215  
Db 2267 TCGCTGTGTGGGAGATAGAGATCTATCATGAGGCGCTCTCTGGGGGCCACCATCTCTGTGATGA 2326



Matches 2179; Conservative 0; Mismatches 1128; Indels 123; Gaps 4;

QY	1002	GCTTGATGGTGTCTCTGCTCCGGTAATAAATCTTTTCTTTGGACATGCAGACAATTCGGG	1061
DB	107	GCTACTTTGAAGACTCTGGATCTGCTGACTTCAGAAGACATTTTGTCAACCTGAGTCCTT	166
QY	1062	AACCGTCAATTTTGTACTGTCTTCATCAAACGATGTGCTGTGATCTGCTCAGATGAGGC	1121
DB	167	CACCATTACTGTGGTCTTACTTCTCAGTGCCTGTGTTTTGTCAACCAAGTTCTCTTTGGAGAAC	226
QY	1122	AGATTTGGAACTGCCACTAGCAGATGGAAGTAACAATTTGTTACGGGAGAGTAGAGGTGAG	1181
DB	227	AGACAGGAGCTGAGGCTAGTGAGTGGTGGAAACAGGTAGCGGGAGAGTGGGAAGTAA	286
QY	1182	AATTCATGAACAGTGGTGGACAATAATGTGCACAGAACCTGGAAGAATGAACAAGCCCTTGT	1241
DB	287	AGTCAGGAGGAGTCGGNAACGGTGTGTAATAATGGCTGGAGCATGGAACGGGTCTCTGT	346
QY	1242	GGTTTGTAAAGCAGTAggATGTCGGTTCAGGCTTTTGGCAGTCTGCTGCTAAACCTAG	1301
DB	347	GATTTGTAAACCAAGCTGGGATGTCCAACCTGCTATCAAGCCCTGGATGGGCTAATTCAG	406
QY	1302	TAAATGAAGCTAGACACATTTGGATAAACAGCATATCTTGCACCTGGGAATGAGTCAGTCT	1361
DB	407	TGCAGGTTCTGGACGCATTTGGATGGATCATGTGTTTGTGCTGGGAATGAGTCAGTCT	466
QY	1362	CTGGGACTGCACATATGATGGAAGAACAAAGC--GAACATGCTTCCGAAGATCAGATGC	1418
DB	467	TTGGGATTTGCAACACATGATGGATGGGGAAGACATAGTAACCTGACTCACCAAGAATGC	526
QY	1419	TGGAGTAAATTTGTTCTGATAAGGCAGATCTGGACCTTAGGCTTGTGGGGCTCATAGCCC	1478
DB	527	TGGAGTGACCTGCTCAGATGGATCCAAATTTGGAATAGGCTGACGCGTGGAGGGAATAT	586
QY	1479	CTGTTATGGGAATTTGGAGGTGAATATACCAAGGAGAGTGGGGGACTGTGTGTCATGACAG	1538
DB	587	GTGTCTGGGAAGTAAGAGATCAAAATCCAAGACGCTGGGGAACAGTGTGTGATGATAA	646
QY	1539	ATGGAGCAAGAAGATCAGAGCTGTGTGTTGTAACAATTTGGCATGTGGAAGCCATATGCA	1598
DB	647	CTTCAACATAGATCATGATCTGTCTATTTGTAGACAACCTTGAATGTGGAAGTCTGTCTAG	706
QY	1599	TGTGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTGGATGACGTTTC	1658
DB	707	TTTCTCTGGTTCTAATAATTTTGGAGAAGGCTCTGGACCAATCTGGTTTGATGATCTTAT	766
QY	1659	TTGCAATTGGAATGAGTCMAATATCTGGGACTGTGGAACAGACTGGATGGGGAAGCATAA	1718
DB	767	ATGCAACGGAAATGAGTCAGCTCTCGAAGCTGCAAGCATCAAGGATGGGGAAGCATAA	826
QY	1719	TTGTGTACACAGAGGATGTGATTGTAACCTGCTCAGGTCATGCAACATGGGGCCCTGAG	1778
DB	827	CTGTGTATCTGTGAGGATGCTGGAGTGATTTGTCTCAAGGGAGCAGATGTAGCCCTTGTAG	886
QY	1779	GCTGTGGGCGGACGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTTCAAGGACGGTG	1838
DB	887	ACTGGTATGATGAGTCACCTGTAATGTTTCAAGGAAGATTAGAAGTGAATTCGAAGGAGATG	946
QY	1839	GGGCACAGTGTGTATGACGGCTGGAAACAGTAAGAGCTGCACCTGTGTGTGTAGCCAGCT	1898
DB	947	GGGGAACAATATGTATGACGGCTGGAGAGTTACGATCTGCTGTGGCATTCGAAGCAACT	1006
QY	1899	GGACTGCCCATCTTCTATCATTTGGCATGGGCTCTGGGAACCTTCTACAGATATGAAA	1958
DB	1007	GGATGTCCAACTCGCGTCAAGCCATTTGGTCGAGTTAACGCCAGTAAGGGATTTTGACA	1066
QY	1959	AATTTTGGCTCGATGATTTTCCCTGTGATGGAGATGAGTCAGATCTCTGGTCATCGAGAA	2018
DB	1067	CATCTGGCTTGACGGGTTTCTTGGCAGGACATGAACCTGCTCTGCGCAATGTAAACA	1126
QY	2019	CAGTGGGTGGGAAATAATGACTGCAGTGCAGTGAAGATGTTGGAGTGATCTCTTCTGA	2078
DB	1127	CCATGAATGGGAAAGCAATATTGCAATTCACAANTGAAGTCTGGCGTGACATCTCTGA	1186

Db	3347	CGGATGCTTTGGGGTGGTTCTGTTGGCCATTTCTGTCATATCTCTTGACTAATAA	3400
Qy	4182	TCAGAAACAAAACATCTCCCTCCTCAGAGTTTCAACACAGAGAGGGGTTCTCTCGAGGA	4241
Db	3407	GCAGAGACAGAGACA--CGCGCTTGCACTTTCCCTCAAGAGAGAGAGAACTTAGTCCACCA	3463
Qy	4242	GAATTTATTTCCATGACATGGAGACCTGCTCAAGAGAGAGAGACCCACATGGGACAAGAC	4301
Db	3464	AAATCAATACCGGAGATGAATTTCTGGCTGAATGCAGATGATCTGGACCTAATGAATTC	3523
Qy	4302	CTCAGATGAC	4311
Db	3524	CTCAGGAGGC	3533
RESULT 11			
ID	AAF45141	AAF45141 standard; cDNA; 4308 BP.	
XX	AC	AAF45141;	
XX	DT	30-MAR-2001 (first entry)	
XX	DE	Bovine WC1 ORF.	
XX	KW	Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;	
XX	KW	central nervous system; focal brain disorder; bipolar affective disorder;	
XX	KW	global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;	
XX	KW	senile dementia; Huntington's disease; amyotrophic lateral sclerosis;	
XX	KW	Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;	
XX	KW	neuropsychiatric; psychoactive substance use; anxiety; ss.	
XX	OS	Bos sp.	
XX	PN	W0200077239-A2.	
XX	PD	21-DEC-2000.	
XX	PF	24-MAY-2000; 2000WO-US14858.	
XX	PR	14-JUN-1999; 99US-0333159.	
XX	FA	(MILL-) MILLENNIUM PHARM INC.	
PI	PI	McCarthy SA, Fraser CC, Sharp JD, Barnes TM;	
XX	DR	WPI; 2001-032313/04.	
XX	DR	P-PSDB; AAB66088.	
XX	PT	TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for	
XX	PT	screening assays and diagnostic assays and for the treatment of	
XX	PT	neurological diseases such as Alzheimer's, Parkinson's and Huntington's	
XX	PT	disease -	
XX	PS	Disclosure; Fig 2; 359pp; English.	
XX	CC	The present invention relates to TANGO or INTERCEPT proteins and codin	
XX	CC	sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057	
XX	CC	AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and codin	
XX	CC	sequences are useful for the treatment of neurological disorders such	
XX	CC	central nervous system (CNS) disorders, CNS-related disorders, focal	
XX	CC	brain disorders, global-diffuse cerebral disorders and other	
XX	CC	neurological and cerebrovascular disorders. The CNS disorders include	
XX	CC	Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic	
XX	CC	lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,	
XX	CC	autonomic function disorders such as hypertension and sleep disorders,	
XX	CC	neuropsychiatric disorders, psychoactive substance use disorders,	
XX	CC	anxiety, and bipolar affective disorder. The present sequence is a	
XX	CC	sequence used in a sequence homology alignment with the TANGO/INTERCE	
XX	CC	sequences of the present invention.	
XX	SQ	Sequence 4308 BP; 866 A; 1093 C; 1410 G; 939 T; 0 other.	

Query Match		19.2%;	Score 886.8;	DB 22;	Length 4308;
Best Local Similarity		54.9%;	Pred. No. 1.8e-214;		
Matches 2061;		Conservative 0;	Mismatches 1527;	Indels 168;	Gaps 9;
Qy	483	TCTGGGTTTGGAGTGTGGATGGAACAACTCTCTGTTTCAGGAGAGAGTGGAGGTGAATTT	542		
Db	75	TCTGGAGCTGAGTGTGAAGGATGAGTCCATCGCTGCTGAGGAGAGTGGAAAGTGAAGCA	134		
Qy	543	CAAAGAAAGTGGGGGACTATATCTGATGATGGTGGAACTTGATCTACTGCTCGCTGGT	602		
Db	135	CAAAGGAGATGGGGCAGTGGATGGTTACAGTGGACATTTGAAGGATGATCTGTAGT	194		
Qy	603	GTGAGGCAACTAGGATGTCCATCTCTCTTTTATTCTTCTGAGTGTGTTAATAGCCCTGC	662		
Db	195	GTGCAGACAGCTGGGGTGTGAGCTGCCATTTGGTTTCTCTGGAGGGCTTATTTTGGGCC	254		
Qy	663	TGTATTGGCCCATTTGGCTGGATGACATTTTATGCCAGGGGAATGATTTGGCACTGTG	722		
Db	255	AGGACTTGGCCCATTTGGCTTTGTATCTATCTGTAAGGGACAGAGTCAACTGTGAC	314		
Qy	723	GAATTGCAGACATCGTGATGGGAAATCATGCTGCACTCACAATAGGATGTCAACATT	782		
Db	315	TGACTGTGAGCATCTAATATTAAGACTATCTGAATGATGGCTATATCAATGTCGGGA	374		
Qy	783	AACTTGTTATGATAGTAGTGTGTAAGTAAAGCTTGTAGGTGGAAGTAAACCCCTGTAT	842		
Db	375	TGCTGGAGTAGTCTGCTCAGGATTGT---TGCGTCTGGCTGGAGGGATGGACCCCTGCTC	431		
Qy	843	GGGAGAGTAGAGTGAATCCAAAGNAAGTGGGGACCCGTATGCCACCAATAGTGAA	902		
Db	432	AGGGCGAGTAGAAGTGCATCTGAGAGAAGTTGGATCCAGTCTGTATGGGAACCTTCAC	491		
Qy	903	CAATGCTGCAGCTGATGCTGATGCAAGCAGTTGGGATGTGGAACCGCACCTTCATCTGC	962		
Db	492	ACTTGGCACTGCCAGATCATCTGTGCAGAGTTGGTGTGGCAGGCTGTGTCTGCTCT	551		
Qy	963	TGGCTTGCCCTCATTTGCAAGTCAGGCTGTGATTTGTATGGTGTGATGGTGTCTCTGCTC	1022		
Db	552	GGGACATGAGCTCTTCAGAGAGTCCAGTGGCCAGGCTCTGGGCTGAAGAGTTTCAGGTGA	611		
Qy	1023	CGGTAATGAATCTTTCTTTGGGACTGCAGACATTCGGGAACCGTCAATTTTGACTGCT	1082		
Db	612	GGGGAGGAGCTGAGCTCTGGGCTGCCCCAGAGTGCCCTGTCCAGGGGGCAGCTGCA	671		
Qy	1083	TCATCAAAACCATGTGCTGTGATCTGCTCAGATGGAGCAGATTGGAACCTGCGACTAGC	1142		
Db	672	CCAGTGGATCTGCTCAGGTTGTTGTTTTCAGCATCTCAGA---AGTCCGGCTCATGAC	728		
Qy	1143	AGATGGAAGTAACAATTTTTCAGGGAGAGTAGAGGTGAGAATTCATGAACAGTGGTGAC	1202		
Db	729	AAACGGCTCCCTCAGTGTGAAGGCGAGGTGGAGATGAACATTTCTGGACAATGGAGAGC	788		
Qy	1203	AATATGTACACAGAACTGGAGATGAACAGCCCTTGTGTTTGTGAAGCAGCTAGGATG	1262		
Db	789	GCTCTGTCCTCCCACTGGAGTGTGGCCAAATGCCAAATTTATCTGTCTCAGCTCGGCTG	848		
Qy	1263	TCCGTTGAGCTCTTTGGCAGTCTGCTGTGCTAAACCTAGTAAATGAAGCTAGAGACATTTG	1322		
Db	849	TGGAGTTGCCATCTCCACCCCGGAGGACCACACTTGGTGAAGAAGGTGATCATGCTT	908		
Qy	1323	GATAAACAGCATATCTTGGCACTGGGAATGAGTCAGCTCTCTGGGACCTGCATATGATGG	1382		
Db	909	AACAGCCGATTTCACTGCTCTGGGCTGAGTCTCTCTCTGGAGTTGCTCTGTGACTGC	968		
Qy	1383	AAAAGCAAGCGACATGCTTCCGAAGATCATGATGCTGGAGTAATTTGTTCTGTATAAGGC	1442		
Db	969	CCTGGGTGGTCTGACTGTTCCCATGGCAACACAGCCCTCTGTGATCTGCTCAGGAAACA	1028		
Qy	1443	AGATCTGG-----	1450		
Db	1029	GATCCAGGTGCTTCCCGAGTGCAACGACTCGGTGTCTCAACCTACAGGCTCTCGCGGCTC	1088		

Qy	1451	-----ACCTAAGGCTTGTGCGGGCTCATAG	1475		
Db	1089	AGAGCACAGCCCCCTACTGCTCAGACAGCAGGCGAGCTCCGCGCTGGTGAGCGGGCGG	1148		
Qy	1476	CCCTGTTATCGGAGATTTGAGGTGAATACCAAGGAGAGTGGGGGACTGTGTGTCATGA	1535		
Db	1149	TCCCTGCGCGGGAGTGGAGATCCTTTGACACAGGGCTCCTGGGCGACCATCTGTGATGA	1208		
Qy	1536	CAGATGAGGACAAAGGAATGAGCTGTTGTGTAAACAATTTGGCATGTGGAACCCCTAT	1595		
Db	1209	CGGCTGGACCTGAGCAGTCCCGGCTGTGTGAGGAGCTGGGCTGTGGAGAGCCCT	1268		
Qy	1596	GCATGTTGTTGATGACCTATTTTAAAGAACATCAGGACCTATTTGGCTGGATGAGCT	1655		
Db	1289	CAATGCCACGGGCTGCTGCTCACTTCGGGGCAGGATCAGGGCCCATCTGTTGGACAACCT	1328		
Qy	1656	TCTTTGCAATTTGGAATGAGTCAATATCTGAGACTGTGAACACAGTGTATGGGAAAGCA	1715		
Db	1329	CAACTGACAGGAAGGAGTCCCAGTGTGAGGTGCCCTTCCCGGGGCTGGGGGACGA	1388		
Qy	1716	TAATTGTGTACACAGAGAGGATGTGATTGAACCTGCTCAGGTGATGCAACATGGGGCT	1775		
Db	1389	CAACTGCAGACACAGCAGGCGGGGTCTATCTGCTCAGAGT-----TCTTGGCCCT	1442		
Qy	1776	AGGCTGTGTGGGCGCAGCAACCGCTGCTCGGGAAGACTGGAGGTGTACTTTCAAGGAGC	1835		
Db	1443	CAGGATGTGAGTAGGAGGACAGCAGTGTCTGGTGGTGGAAAGTTTCTTACAATGGGAC	1502		
Qy	1836	CTGGGCACTGTGTGATGAGCGGTGGAACAGTAAAGCTGCAGCTGTGGTGTGTAGCA	1895		
Db	1503	CTGGGCGAGTGTCTGCCGTAAACCCCATGGAACATCACTGTGTCCACGATCTGCAGACA	1562		
Qy	1896	GCTGACGTGCCCATCTTCTATATTGGCATGGGTCTGGGAAACGCTTCTACAGGATATGG	1955		
Db	1563	CCTTGGCTGTGGGACAGTGAACCCCTCACTCTCTTCTGTCTTTAGAGAAGTTTAG	1622		
Qy	1956	AAAAATTTGGCTCGATGATGTTTCTGTGATGGAGATGAGTCACTCTCTGCTATGATGAG	2015		
Db	1623	GCCACAGTGGTGGATAGATCCAGTGTGGGAAACTGACACCTCTCTCTGSCAGTGCC	1682		
Qy	2016	GACAGTGGTGGGGAATTAATGACTGCAGTGCACAGTGAAGATGTTGGAGTGTCTGTC	2075		
Db	1693	TCTGACCTTTGGAATTACAACCTCATGCTCTCCAAGGAGGAGCCTATATCTGTGTG---	1739		
Qy	2076	TGATGCTCGGATATGAGGCTGAGGCTTGTGGTGTGGAAGCAGCAGCTGTGCTGGAAGCT	2135		
Db	1740	---TGCACACAGCAGACAGATCCCGCTGTGGATGGAGGTGTCTCTCTGCGAGAGT	1796		
Qy	2136	TGAGTGAATGTCCAGGTGCGGTGGGAATTTCTGTGCTATATGCTGGGGAATGAACAT	2195		
Db	1797	GGAGATCCTTGACCAAGGCTCTCTGGGCGACCATCTGTGATGACCGCTGGACCTGGACGA	1856		
Qy	2196	TGCTCAAGTTGTTCCAGGCAACTGAATGTGGGTCTGCAATCAGGGTCTCCAGAGAGCC	2255		
Db	1857	TGCCCGTGTGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTGGAGCCACTGTCTCTTC	1916		
Qy	2256	TCAATTCACAGAAACATTAACATCTTAATGTGCAATTTCTGCTGCTCACTGGAGGGA	2315		
Db	1917	CTTCTTCGGGAGGATCAGGGCCCTCTGGCTGGATGAGTGAAGTGAAGTGAAGGAGAGA	1976		
Qy	2316	AGCCTCTCTCTGGGATTTATACGATGGAGTGGAAACAGACTGCGTGTCTATTTAAATAT	2375		
Db	1977	GTCCCAAGTATGGAGTGCCTCTCTGGGATGGGGCAACACAACTGCAATCATCAAGA	2036		
Qy	2376	GGAGCAAGTTTGTCTCTCAGCCACAGGAGCCCAAGCTGGTGTGGAGTGTATGCC	2435		
Db	2037	AGATGCAAGGATCTCTCTCAGGATTTGTGC-----GTCTGGCTGGAGGAGATGGACC	2090		
Qy	2436	CTGCTCTGAGCTGTGAGTGAACATGCACACATGCGGCTCTGTCTGTGATTTCTGA	2495		
Db	2091	CTGCTCAGGGGAGTAGAAGTGCATTTGGAGAAGCCCTGGACCCCAAGTGTCTGATGGAAA	2150		
Qy	2496	TTTCTCTCTTCAATGCTGCCAATGTGCTGTGCAGAGAATTAATTTGGAGATGCCATATC	2555		

Db 2151 CTTTACACCTCCACAGTCCAGGTCATCTGTGAGAGCTGGGATGTGGGAAGGCTGTGC 2210  
QY TCTTTCTGTGGGAGATCACTTTTGGAAAGGGAATGCTTAACCTTGGGCCGCGAAAGTTCCA 2615  
Db 2211 TGTCTGGGACATGCCATTAGAGAGTCGATGCCAGGCTGGGCTGAAGAGTTGAG 2270  
QY GTGTGAAGGAGTGAACATCACTTGCATATATGCCCCTATGTTCAACATCCGCGAAGACAC 2675  
Db 2271 GTGTGATGGGGGAGGCTGAGCTGTGGTCTGCCACAGATGCCCTGTGCCAGGAGCAC 2330  
QY TTTCTATCCACAGACAGAGAGTTGGAGTGTCTGTCTTCCCGATATACAGATGTCCGACTTGT 2735  
Db 2331 ATCTCTCCACAGTGGAGTGTCTGAGTGTCTGTCTGAGTGTACACAGAGATCCAGCTTAT 2390  
QY G---AATGCGAAATCCAGTGCACGGGCAAGTGGAGATCAAGCTGCTTGGACACTGGGG 2792  
Db 2391 GAAAACGGCACCCTCAATGTGAGGGCAGTGGAGATGAAGATCTCTGGAGATGGAG 2450  
QY CTCACCTGTGTGACACCCACTGGGACCCAGAGATGCCCGTGTCTATGAGACAGACTCAG 2852  
Db 2451 AGCGCTGTGTCTCCACTGTGAGTGTGGCCAAATGCAATGTGTCTGTCTGCTCAGCTCG 2510  
QY CTGTGGGACTGTCTCTCAACACAGAGAGAAATATATTGGAGAAAGTGTCTGTCTGT 2912  
Db 2511 CTGTGAGTGGCCTCTCCACCCAGAGACCACTTGGTGGAGGAGGTGATCAGAT 2570  
QY GTGGGACACAGGTTTCAATGTCTAGGGAATGAGTCACTTCTGGATAAATCTCAAAATGAC 2972  
Db 2571 CTCAACAGGCCAATTTCACTGTCTAGGGCTGAGTCTCTCTGAGTGTCTCTGTGAC 2630  
QY AGTCTTGTGAGCAGCTCCCTGTATTCATGGAATATCTGTCTGTATCTGTATCTGACAGAG 3032  
Db 2631 TGCTTTGGGTGGGCTGACTGTTCCTACTGCTCAGACAGCAGGCTCTGTATCTGTCTGAG 2690  
QY CTGTGACCCAGCCTGTCTTCCATGCTCTGCAATGATCTGACCCCAATTTGTCTGCTCAGT 3092  
Db 2691 CCACCCAGGTCCTGCCCCAGTGCACAGCTTCTGTCTCAACCTGACGGCTCTCGGC 2750  
QY TCCAGAGGCGAGTCTTGTGATCTGTAGGAGCAACAGGCTCCGCTAGTGGATGGGA 3152  
Db 2751 CTCAGAGGAGTCTTCTCCCTACTGCTCAGACAGCAGGAGCTCCGCTGTGTGACGGGG 2810  
QY CAGCCGCTGTGCGGAGAGTAGAGATCTATCAGAGCCTCTGTGGGACCACTCTGTGA 3212  
Db 2811 CGTCTCCCTGCGGCGGAGAGTGGAGATCTTGACAGGCTCTCTGGGACCACTCTGTGA 2870  
QY TGACGGCTGGGACCTGAGCGATGCCACAGTGGTGTGTCAAAAGCTGGGCTGTGGAGTGGC 3272  
Db 2871 TGATGACTGGGACCTGGACGATGCCGCTGTGTGTGTCAGGACGCTGGGCTGTGGAAAGC 2930  
QY CTTCAATGCCAGGCTCTGTCTCACTTTTGGGAGGGGTGAGGGCCCATCTGGCTTGGATGA 3332  
Db 2931 CCTCAATGCCAGGGCTGTCTCACTTCGGGGCAGGATCAGGGCCCATCTGGCTGACGA 2990  
QY CTTGAATGCCAGAGAACGAGTCCACTTGTGGAGTGGCTTCCCGGGCTGGGGGCA 3392  
Db 2991 CTTGAATGCCAGAGAAAGAGTCCACAGTGTGGAGTGGCTTCCCGGGCTGGGGGG 3050  
QY GCACGACTGCAGGACAGAGGAGCAGGGGTCACTGTCTCAGAAATTCACAGCCTTGAG 3452  
Db 3051 GCACGACTGCAGACACAGGAGAGCCCGGGTCACTGTCTCAGAGTCTCTGGCCCTCAG 3110  
QY GCTCTACGTGAACATGAACAGAGAGCTGTGTCTGGGAGATTGGAAGTCTTCTATACCGG 3512  
Db 3111 G-----ATGGTCAGCGAGGACCAAGCAGTGTGTGGTGGCTGGAGGTTTCTTACAACGG 3164  
QY GACCTGGGCGAGCTGGCAGGAGGAGCAATCACCAGCCATAGCAGGATGTGTGCGAG 3572  
Db 3165 GACCTGGGCGAGTGTGTGGCGAGCCCAATGGAAGATATACGTGTCTGCTGATCTGCGAG 3224  
QY GCAGCTGGGCTGTGGGAGAAATGGAGTGTGTGACGCTTCGCCCTTTATCTAAGACAGGCTC 3632

Db 3225 ACAGCTTGGATGTGGGGACAGTGGAAAGTCTCAACACCTCTGTGGTCTCAGGAAAGTTTC 3284  
QY TGGTTTTCATGTGGGTGGATGACATTCAGTGTCTCTAAAGCATATATCTCCATATGCGAGTG 3692  
Db 3285 TAGACCCCGGTGGGTAGATTAAATTCAGTGTGGAAATGGATACCTCTCTCTGGCAGTG 3344  
QY CTTGTCTGCCATCGGAGGCGGAAGATCTCCAGCCACAGCAAGAGACCTCTGGATCAGTG 3752  
Db 3345 TCTCTTGTGCCATCGAAATACAGTTTCATGCTCTCCAAAGGAGAGCCTTACATCTCATG 3404  
QY TGAAGATAGAATAA----- 3766  
Db 3405 TGAAGGAAGAAGACCAAGAGCTGTCCAAGTGTCCGCGCTGCACAGACAGAGAGAGCT 3464  
QY --GAGTGCCTGGAGGAGACACCGAGTGTCTGGGAGAGTGGAGATCTGGCAGCAGCGCTC 3824  
Db 3465 CCGCCTCAGGGGAGGAGACAGCGAGTGTCTCAGGGCGGTGGAGGTGTGGCACAACGCTC 3524  
QY CTGGGGCAGAGTGTGTGATGACTCTCTGGACCTTGGCCGAGGCGGAAGTGTGTGTGAGCA 3884  
Db 3525 CTGGGGCAGCGTGTGCGATGACTCTCTGGAGCCTTGGCAGAGGCTGAGGTGTGTGTGAGCA 3584  
QY GCTGGGCTGTGGCTGTCTGCTGCTGCTGAGGGAGCGTTCGTTTGGCCAGGAACTGG 3944  
Db 3585 GCTGGGCTGTGGCCAGCGCTTGAAGCGTGTGCGGTCTGAGCAATTTGGCCCTGGAATGG 3644  
QY AACCATCTGTGGTGTGATGACATCGGTGCAAGAGAAATAGTCAATTTCTATGGGACTGTCA 4004  
Db 3645 GAGCATCTGGTGGACGAGTGTGAGTGTGGCGGGCGGGAGTCTCTCTGTGGGACTGTGT 3704  
QY CGCCAAACCTTGGGACAGAGTGTGTGACACAGGAGAGATGCTGGGCTGAGGTGCTC 4064  
Db 3705 TCGGAGCGCTTGGGCGAGCGCACTGCAAGCAGCAGGAGGATGCTGGTGTGAGGTGCTC 3764  
QY TGGACAGTGTGCTGAAATCACTGAATGCTCTCTCAGG 4100  
Db 3765 TGGTGTGAAGACAAACATTGCCACGACACAGCAGG 3800

RESULT 12  
AAAA3017  
ID AAA43017 standard; cDNA; 820 BP.  
XX  
AC AAA43017;  
XX  
DT 21-AUG-2000 (first entry)  
XX  
Human secreted expressed sequence tag SEQ ID NO:1757.  
DE  
XX  
KW Human; mouse; xenopus; rat; secreted expressed sequence tag; EST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;  
KW antiviral; antidiabetic; antiasthmatic; vulnerary; antiparkinsonian;  
KW antiulcer; osteopathic; neuroprotective; nootropic; antipsoriatic;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
OS  
XX Homo sapiens..  
XX  
PN WO200021990-A1.  
XX  
PD 20-APR-2000.  
XX  
PF 15-OCT-1999; 99WO-US24205.  
XX  
PR 15-OCT-1998; 98US-0104435.



PT Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection  
PS Claim 1; Page 341; 1217pp; English.  
XX  
XX  
XX AAH99166 to AAH99904 encode the human proteins given in AAH25225 to  
CC AAH25963. The proteins can have activities based on the tissues and  
CC cells they are expressed in, such as: antinflammatory; antirheumatic;  
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnerary;  
CC antidiabetic; osteopathic; dermatological; antiallergic; antiasthmatic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX  
XX  
SQ Sequence 690 BP; 163 A; 134 C; 207 G; 186 T; 0 other;  
Query Match 13.9%; Score 642; DB 22; Length 690;  
Best Local Similarity 99.2%; Pred. No. 1.1e-152;  
Matches 645; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1930 CTGGAAACGCTCTACAGGATATGGAATAATTTGGTCGATGATTTCTCTGTGATGGA 1989  
DB 40 CGGGAAACGCTCTACAGGATATGGAATAATTTGGTCGATGATTTCTCTGTGATGGA 99  
QY 1990 GATGAGTCAGATCTCTGTCATGAGCAAGTGGGGTGGGAAATAATGACTGCAGTCAC 2049  
DB 100 GATGAGTCAGATCTCTGTCATGAGCAAGTGGGGTGGGAAATAATGACTGCAGTCAC 159  
QY 2050 AGTGAAGATCTGGAGTATCTGTCATGATGATGATGATGATGATGATGATGATGATG 2109  
DB 160 AGTGAAGATCTGGAGTATCTGTCATGATGATGATGATGATGATGATGATGATGATG 219  
QY 2110 GGAAGCAGCAGGTGTCGTCGAAAGTGTAGGTGAATGTCAGGCGTCCGCGGAATTCGT 2169  
DB 220 GGAAGCAGCAGGTGTCGTCGAAAGTGTAGGTGAATGTCAGGCGTCCGCGGAATTCGT 279  
QY 2170 TGTGCTAATGGTGGGAATGACATTCGTCGAGTGTGTCGAGGCAACTGTAATGTGGG 2229  
DB 280 TGTGCTAATGGTGGGAATGACATTCGTCGAGTGTGTCGAGGCAACTGTAATGTGGG 339  
QY 2230 TCTGCAATCAGGCTCTCCAGAGCCCTCATTTCACAGAAAGACATTCACATCTTAATG 2289  
DB 340 TCTGCAATCAGGCTCTCCAGAGCCCTCATTTCACAGAAAGACATTCACATCTTAATG 399  
QY 2290 TCGAATTTCTGGCTGCACTGGAGGGAAAGCCTCTCTCTGGGATGTATACATGGGAGTGG 2349  
DB 400 TCGAATTTCTGGCTGCACTGGAGGGAAAGCCTCTCTCTGGGATGTATACATGGGAGTGG 459  
QY 2350 AACACAGACTCGTGTCTAATTAATATGGAAGCAAGTTTGTATCTGCTCAGGCCACAGGCAG 2409  
DB 460 AACACAGACTCGTGTCTAATTAATATGGAAGCAAGTTTGTATCTGCTCAGGCCACAGGCAG 519  
QY 2410 CCCAGGCTGGTGGAGTGATATCCCTGCTCTGGAGCTGTGAAGTGAACATGCGAGAC 2469  
DB 520 CCCAGGCTGGTGGAGTGATATCCCTGCTCTGGAGCTGTGAAGTGAACATGCGAGAC 579  
QY 2470 ACATGGCGCTCTGCTCTGATTTCTGATTTCTCTCTCATCTGCGCAATGTGCTGTGAGA 2529  
DB 580 ACATGGCGCTCTGCTCTGATTTCTGATTTCTCTCTCATCTGCGCAATGTGCTGTGAGA 639

QY 2530 GAATTAATTTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGG 2579  
DB 640 GAATTAATTTGGAGATGCCATATCTCTTTCTGTGGGAGATCACTTTGG 689  
RESULT 14  
AAV89272  
ID AAV89272 standard; cDNA; 608 BP.  
XX  
XX AAV89272;  
XX  
XX 15-FEB-1999 (first entry)  
XX EST clone CF118.  
XX Human; secreted protein; expressed sequence tag; EST; haematopoiesis;  
KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;  
KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;  
KW gene therapy; ss.  
XX Homo sapiens.  
XX WO9845436-A2.  
XX  
XX 15-OCT-1998.  
XX  
XX 10-APR-1998; 98WO-US06955.  
XX  
XX 10-APR-1997; 97US-0838821.  
XX (GEMY) GENETICS INST INC.  
XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
PI Racie LA, Spaulding V, Treacy M;  
XX WPI; 1999-070077/06.  
XX  
XX New polynucleotides encoding human secreted proteins - derived from  
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
PT ovary, pituitary, retina and colon cDNA libraries.  
XX  
XX Claim 1; Page 165; 618pp; English.  
XX  
XX The present sequence represents a human expressed sequence tag (EST).  
CC The polynucleotide, which is a secreted EST, and the encoded protein  
CC are predicted to have useful biological activities which would make  
CC them suitable for treating, preventing or ameliorating medical  
CC conditions in humans and animals, although no supporting data is  
CC given. Suggested activities include nutritional activity, immune  
CC stimulating or suppressing activity, haematopoiesis regulating  
CC activity, tissue growth activity, activin/inhibin activity,  
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic  
CC activity, receptor/ligand activity, anti-inflammatory activity,  
CC cadherin/tumour invasion suppressor activity, tumour inhibition  
CC activity. The polynucleotide may also be useful for gene therapy.  
XX  
XX  
SQ Sequence 608 BP; 145 A; 109 C; 176 G; 178 T; 0 other;  
Query Match 12.2%; Score 563; DB 20; Length 608;  
Best Local Similarity 99.8%; Pred. No. 1.2e-132;  
Matches 574; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 26 TAATGATGCTGCCCTCAAACTCGTGGCATATTTGATTTGGGAAGATGCTGTCATCAGA 85  
DB 24 TAATGATGCTGCCCTCAAACTCGTGGCATATTTGATTTGGGAAGATGCTGTCATCAGA 83  
QY 86 ACCTTTTCTGCTGTGTTAACTTGCATCTGCTCCCTGAATTCCTTCTCTCATCAGCA 145  
DB 84 ACCTTTTCTGCTGTGTTAACTTGCATCTGCTCCCTGAATTCCTTCTCTCATCAGCA 143  
QY 146 GTTTTAATGAACAGATTTGAGTTGAGGCTTGGCTTGAATGGAGACGGTCCCTGCTCTGGA 205  
DB 144 GTTTTAATGAACAGATTTGAGTTGAGGCTTGGCTTGAATGGAGACGGTCCCTGCTCTGGA 203

QY 206 CAGTGGAGGTGAAATTCAGGACAGTGGGGACTGTGTGATGATGGGTGGAACTA 265  
|||||  
Db 204 CAGTGGAGGTGAAATTCAGGACAGTGGGGACTGTGTGATGATGGGTGGAACTA 263  
|||||  
QY 266 CTGGCTCAACTGTCGTGTGCAACAGCTGGATGCCA-TTTTCTTTGGCCATGTTTGGT 324  
|||||  
Db 264 CTGGCTCAACTGTCGTGTGCAACAGCTGGATGCCA-TTTTCTTTGGCCATGTTTGGT 323  
|||||  
QY 325 TTTGGACAAGCCGTGACTAGACATGGAATAATTTGGCTTGATGTTTCTCTGTATGGA 384  
|||||  
Db 324 TTTGGACAAGCCGTGACTAGACATGGAATAATTTGGCTTGATGTTTCTCTGTATGGA 383  
|||||  
QY 385 AATGAGTCAGCTCTCTGGGAATGCAACACCGGAATGGGAAGCCATAACTGTTATCAT 444  
|||||  
Db 384 AATGAGTCAGCTCTCTGGGAATGCAACACCGGAATGGGAAGCCATAACTGTTATCAT 443  
|||||  
QY 445 GGAGAAGATGTTGGTGAACCTGTTATGGTGAAGCCAACTCGGTTGAGGCTAGTGAT 504  
|||||  
Db 444 GGAGAAGATGTTGGTGAACCTGTTATGGTGAAGCCAACTCGGTTGAGGCTAGTGAT 503  
|||||  
QY 505 GGAACAACTCTCTTTCAGGAGAGTGGAGTGAATTCRAAGAAAGTGGGGACTATA 564  
|||||  
Db 504 GGAACAACTCTCTTTCAGGAGAGTGGAGTGAATTCRAAGAAAGTGGGGACTATA 563  
|||||  
QY 565 TGTGATGATGGGTGGAATTCGAATGCTGCTGCCGT 599  
|||||  
Db 564 TGTGATGATGGGTGGAATTCGAATGCTGCTGCCGT 598  
|||||

RESULT 15  
AAH99326  
ID AAH99326 standard; cdna; 562 BP.  
XX  
AC AAH99326;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Human protein encoding cdna sequence SEQ ID NO:161.  
XX  
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiaggregant; haemostatic; vulnerary; antilulcer; osteopathic; eczema;  
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153455-A2.  
XX  
PD 26-JUL-2001.  
XX  
PF 22-DEC-2000; 2000WO-US35017.  
XX  
PR 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-457603/49.

P-PSDB; AAM25385.

Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection

Claim 1: Page 369; 1217pp; English.

AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary; antidiabetic; osteopathic; dermatological; antiallergic; antiasthmatic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

Sequence 562 BP; 125 A; 131 C; 189 G; 117 T; 0 other;

Query Match 11.9%; Score 551.4; DB 22; Length 562;  
Best Local Similarity 99.8%; Pred. No. 1e-129;  
Matches 552; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3485 CTGGGAGATTGGAAGTCTTCTATACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCA 3544  
|  
Db 10 CCGGAGATTGGAAGTCTTCTATACGGGACCTGGGGCAGCGTCGGCAGGAGGAACATCA 69  
|  
QY 3545 CCACAGCATAGCAGGACCTGTCAGCAGCTGGGCTGTGGGAGAGTGGAGTTGCA 3604  
|  
Db 70 CCACAGCATAGCAGGACCTGTCAGCAGCTGGGCTGTGGGAGAGTGGAGTTGCA 129  
|  
QY 3605 GCCTCGCCCTTTATCTAAGACAGGCTCTGGTTTCATCTGGGTGGATGACATTCAGTCTC 3664  
|  
Db 130 GCCTCGCCCTTTATCTAAGACAGGCTCTGGTTTCATCTGGGTGGATGACATTCAGTCTC 189  
|  
QY 3665 CTAAACCCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGATCTCCA 3724  
|  
Db 190 CTAAACCCATATCTCCATATGGCAGTGCCTGTCTGCCCCATGGGAGCGAAGATCTCCA 249  
|  
QY 3725 GCCCAGCAGAGAGACCTGGATCAGATCTGAAGATAGATAAGAGTGGCTGGAGGAGACA 3784  
|  
Db 250 GCCCAGCAGAGAGACCTGGATCAGATCTGAAGATAGATAAGAGTGGCTGGAGGAGACA 309  
|  
QY 3785 CCGAGTGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGGGCAGTGTGTGATG 3844  
|  
Db 310 CCGAGTGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGGGCAGTGTGTGATG 369  
|  
QY 3845 ACTCTGGGACCTGGCCGAGGCGGAAGTGTGTGTCACAGCTGGGCTGTGCTCTGCTC 3904  
|  
Db 370 ACTCTGGGACCTGGCCGAGGCGGAAGTGTGTGTCACAGCTGGGCTGTGCTCTGCTC 429  
|  
QY 3905 TGGCTGCCCTGAGGACGCTTCGTTGGCCAGGGAACCTGGAACCATCTGGTGGATGACA 3964  
|  
Db 430 TGGCTGCCCTGAGGACGCTTCGTTGGCCAGGGAACCTGGAACCATCTGGTGGATGACA 489  
|  
QY 3965 TCCGGTGC AAAAGAAATGAGTCATTTCTATGGGACTGTACGCCCAACCCCTGGGACAGA 4024  
|  
Db 490 TCCGGTGC AAAAGAAATGAGTCATTTCTATGGGACTGTACGCCCAACCCCTGGGACAGA 549  
|  
QY 4025 CTGACTGTGGACA 4037  
|

Db 550 GTGACTGTGGACA 562

Search completed: May 12, 2003, 03:08:03  
Job time : 683 secs

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 03:08:10 ; Search time 50 Seconds  
(without alignments)  
3872.263 Million cell updates/sec

Title: US-09-759-130B-381  
Perfect score: 8138  
Sequence: 1 MMLPQNSWHIDFGRCCHQN.....CEDASDTSLGLVLPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8138	100.0	1453	22	Human TANGO 234.
2	7913	97.2	1413	22	Human TANGO 234 ma
3	7426	91.3	1319	22	Human TANGO 234 ex
4	6058	74.4	1120	22	Human secreted pro
5	4540	55.8	821	22	Human full-length
6	3441.5	42.3	1151	23	Human CD163 recept
7	3429	42.1	1121	22	Human polypeptide
8	3429	42.1	1124	22	Human polypeptide
9	3429	42.1	1124	22	Human polypeptide
10	3429	42.1	1156	23	Human CD163 recept

11	3424	42.1	1116	23	AAU97582	Human CD163 recept
12	3396.5	41.7	1149	22	AAU97583	Human CD163 recept
13	3396.5	41.7	1154	22	AAU97583	Human polypeptide
14	3082	37.9	1436	22	AAU97583	Bovine WCL protein
15	2435	29.9	1785	19	AAU97583	Human SRCR protein
16	1763.5	21.7	898	13	AAU97583	Novel human secret
17	1183	14.5	225	22	AAU97583	Human protein sequ
18	1048.5	12.9	1290	18	AAU97583	Rat von Ebner's gl
19	1033	12.7	186	22	AAU97583	Human protein sequ
20	971.5	11.9	552	22	AAU97583	Human sbg14862SPER
21	921	11.3	422	22	AAU97583	Human sbg14862SPER
22	822.5	10.1	822	20	AAU97583	Human serine prote
23	822.5	10.1	875	20	AAU97583	Human neurotrophin
24	753	9.3	143	22	AAU97583	Human contig polyp
25	725	8.9	141	22	AAU97583	Human M160 precurs
26	719.5	8.8	761	20	AAU97583	Mouse serine prote
27	719.5	8.8	761	20	AAU97583	Mouse neurotrophin
28	712.5	8.8	666	19	AAU97583	Human SRCR protein
29	691.5	8.5	347	19	AAU97583	Amino acid sequenc
30	691.5	8.5	347	20	AAU97583	Human PRO229 prote
31	691.5	8.5	347	22	AAU97583	Human angiogenesis
32	691.5	8.5	347	23	AAU97583	Human PRO229 prote
33	691.5	8.5	347	23	AAU97583	Human PRO229 prote
34	691.5	8.5	347	23	AAU97583	Human liver cell c
35	691.5	8.5	347	19	AAU97583	Polypeptide isolat
36	634.5	7.8	757	21	AAU97583	Human Cys-rich sca
37	634	7.8	111	22	AAU97583	Human lysyl-oxidas
38	633.5	7.8	757	23	AAU97583	Human lysyl-oxidas
39	632.5	7.8	732	22	AAU97583	Human CG153 (Or C5
40	629	7.7	754	21	AAU97583	Murine lysyl oxida
41	627.5	7.7	753	22	AAU97583	Human CG153 (Or C5
42	625.5	7.7	743	23	AAU97583	Human lysyl oxidas
43	624.5	7.7	641	21	AAU97583	Human secreted pro
44	622.5	7.6	752	23	AAU97583	Human lysyl-oxidas
45	622.5	7.6	753	21	AAU97583	Human lysyl oxidas

ALIGNMENTS

RESULT 1

AAU97582  
ID AAB66037 standard; Protein; 1453 AA.

XX AAB66037;

AC AAB66037;

DT 30-MAR-2001 (first entry)

XX Human TANGO 234.

DE TANGO protein; INTERCEPT protein; neurological disorder;  
KW central nervous system; focal brain disorder; bipolar affective disorder;  
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;  
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;  
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;  
KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

PN WO200077239-A2.

XX 21-DEC-2000.

PD 24-MAY-2000; 2000WO-US14858.

PF 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

DR N-PSDB; AAF45123, AAF45124.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for  
PT screening assays and diagnostic assays and for the treatment of  
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's  
PT disease -

XX Claim 8; Fig 2; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding  
XX sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,  
XX AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding  
XX sequences are useful for the treatment of neurological disorders such as  
XX central nervous system (CNS) disorders, CNS-related disorders, focal  
XX brain disorders, global-diffuse cerebral disorders and other  
XX neurological and cerebrovascular disorders. The CNS disorders include  
XX Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic  
XX lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,  
XX autonomic function disorders such as hypertension and sleep disorders,  
XX neuropsychiatric disorders, psychoactive substance use disorders,  
XX anxiety, and bipolar affective disorder.

XX Sequence 1453 AA;

Query Match 100.0%; Score 8138; DB 22; Length 1453;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMLPONSWHIDGRCCCHONLFSAVVTCILLNSCLFSSNGTDLRLVNGDPCSGT 60  
DB 1 MMLPONSWHIDGRCCCHONLFSAVVTCILLNSCLFSSNGTDLRLVNGDPCSGT 60  
QY 61 VEVKQGGWGTVCDDGWNTTASTVYCKQLGCPFPFAMRFGQAVTRHGKIWLDDVSCYGN 120  
DB 61 VEVKQGGWGTVCDDGWNTTASTVYCKQLGCPFPFAMRFGQAVTRHGKIWLDDVSCYGN 120  
QY 121 ESALMEQHQREWSHNCYHGEDVGYNCYCEANLGLRLVDGNNSCSGRVEVRFQERWGTIC 180  
DB 121 ESALMEQHQREWSHNCYHGEDVGYNCYCEANLGLRLVDGNNSCSGRVEVRFQERWGTIC 180  
QY 181 DQGNLNTAAVYCRQLGCPSPFISGVVNSPAVLRLPWLDDTLCOGNELALWNCRRHGWG 240  
DB 181 DQGNLNTAAVYCRQLGCPSPFISGVVNSPAVLRLPWLDDTLCOGNELALWNCRRHGWG 240  
QY 241 NHDCSHNEDVTLCYDSSDLELRLVGGTNRGMRVELKIQGRWGTVCCHHKWNNAAADVVC 300  
DB 241 NHDCSHNEDVTLCYDSSDLELRLVGGTNRGMRVELKIQGRWGTVCCHHKWNNAAADVVC 300  
QY 301 KOLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFELWDCRHSCTVNFDCLEHNDVSI 360  
DB 301 KOLGCGTALHFAGLPHLOSGSDVWLDGVSCGNSFELWDCRHSCTVNFDCLEHNDVSI 360  
QY 361 CSDGADLELRLADGNNSCSGRVEVRIHQWWTICDQNKNEQALVVCQLGCPFSVFGSR 420  
DB 361 CSDGADLELRLADGNNSCSGRVEVRIHQWWTICDQNKNEQALVVCQLGCPFSVFGSR 420  
QY 421 RAKPSNEARDIWNISCTGNESALWDCYTDGKARTCFRRSDAGVICSDDKADLRLVG 480  
DB 421 RAKPSNEARDIWNISCTGNESALWDCYTDGKARTCFRRSDAGVICSDDKADLRLVG 480  
QY 481 AHSPPCYGRLEVKYQGEWGTVCCHDRSTRNAAVVCQLGCGKPMHVFGMTYFKAESGPIWL 540  
DB 481 AHSPPCYGRLEVKYQGEWGTVCCHDRSTRNAAVVCQLGCGKPMHVFGMTYFKAESGPIWL 540  
QY 541 DDVSCIGNESNIWDCHEHSGWGHKNCVHREDVITVCSGATWGLRLVGGNRCSGRLEVYF 600  
DB 541 DDVSCIGNESNIWDCHEHSGWGHKNCVHREDVITVCSGATWGLRLVGGNRCSGRLEVYF 600  
QY 601 QGRWGTVCDDGNNKNSAAAVVCSQLDPCPSIIICMGLGNASTGYGKIWLDDVSCOGDESILW 660  
DB 601 QGRWGTVCDDGNNKNSAAAVVCSQLDPCPSIIICMGLGNASTGYGKIWLDDVSCOGDESILW 660  
QY 661 SCRNMGWNNDCSHSDVGVICSDADMELRLVGGSSRCAGKVEVNVQAVGILCANGWG 720

DB 661 SCRNMGWNNDCSHSDVGVICSDADMELRLVGGSSRCAGKVEVNVQAVGILCANGWG 720  
QY 721 MNIAEVVCRQLECCSATRVSRPHFTERTLHILMNSNGCTGEASLWDCIRWENKQTACH 780  
DB 721 MNIAEVVCRQLECCSATRVSRPHFTERTLHILMNSNGCTGEASLWDCIRWENKQTACH 780  
QY 781 LNMEASLICSAAHROPRLVADMPGSGRVEVKHADTWRSVCDSDFSLSHAANVLCRELNGCD 840  
DB 781 LNMEASLICSAAHROPRLVADMPGSGRVEVKHADTWRSVCDSDFSLSHAANVLCRELNGCD 840  
QY 841 AISLSVGDHFGKNGLTWAERFQCEGSETHALCPVIOHPEDTCHISREVGVVCSRYTDV 900  
DB 841 AISLSVGDHFGKNGLTWAERFQCEGSETHALCPVIOHPEDTCHISREVGVVCSRYTDV 900  
QY 901 RLNVKSGOCDQOVENLVGHWSGLCDTHWDPEDARVLCQLSCGTALSTTGKYGIGERSV 960  
DB 901 RLNVKSGOCDQOVENLVGHWSGLCDTHWDPEDARVLCQLSCGTALSTTGKYGIGERSV 960  
QY 961 RVWGRHFRHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLS 1020  
DB 961 RVWGRHFRHCLGNESLLDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLS 1020  
QY 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEYHDFGFWGTICDDGMDLSDAHVVCOKLGG 1080  
DB 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEYHDFGFWGTICDDGMDLSDAHVVCOKLGG 1080  
QY 1081 VAFNATVSAHFEGEGSGPIWLDLNCCTGTESHLMWQCPSRGWGQHDCHRHKEDAGVICSEFTA 1140  
DB 1081 VAFNATVSAHFEGEGSGPIWLDLNCCTGTESHLMWQCPSRGWGQHDCHRHKEDAGVICSEFTA 1140  
QY 1141 LRLYSETETESCAGRLEVYNGTWGSGRRNITTAIAGIVCRQLCGGNGVVSAPLSKT 1200  
DB 1141 LRLYSETETESCAGRLEVYNGTWGSGRRNITTAIAGIVCRQLCGGNGVVSAPLSKT 1200  
QY 1201 GSGFMWVDIOCPKTHISIWQCLSAWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260  
DB 1201 GSGFMWVDIOCPKTHISIWQCLSAWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260  
QY 1261 IWHAGSWGTVCCDSDWDLAEAEVVCQOLGCGSALALRDASFQGGTGTIWLDDMRCKGNES 1320  
DB 1261 IWHAGSWGTVCCDSDWDLAEAEVVCQOLGCGSALALRDASFQGGTGTIWLDDMRCKGNES 1320  
QY 1321 FLWDCHAPWGSDCHGHKEDAGVRCGSLKSLNASSGHLALILSSIFGLLLVLFLFL 1380  
DB 1321 FLWDCHAPWGSDCHGHKEDAGVRCGSLKSLNASSGHLALILSSIFGLLLVLFLFL 1380  
QY 1381 TWCVRQKOKHLPRLVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDTTPNHGCCDASDT 1440  
DB 1381 TWCVRQKOKHLPRLVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDTTPNHGCCDASDT 1440  
QY 1441 SLLGVLPASEATK 1453  
DB 1441 SLLGVLPASEATK 1453  
RESULT 2  
AAB66039  
ID AAB66039 standard; Protein; 1413 AA.  
XX  
AC AAB66039;  
XX  
DT 30-MAR-2001 (first entry)  
XX  
DE Human TANGO 234 mature protein.  
XX  
KW TANGO protein; INTERCEPT protein; neurological disorder;  
KW central nervous system; focal brain disorder; bipolar affective disorder;  
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;  
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;  
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;  
KW neuropsychiatric; psychoactive substance use; anxiety.  
OS Homo sapiens.



Human TANGO 234 extracellular domain.

TANGO protein; INTERCEPT protein; neurological disorder;  
central nervous system; focal brain disorder; bipolar affective disorder;  
global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;  
senile dementia; Huntington's disease; amyotrophic lateral sclerosis;  
Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;  
neuropsychiatric; psychoactive substance use; anxiety.

Homo sapiens.

WO200077239-A2.

21-DEC-2000.

24-MAY-2000; 2000WO-US14858.

14-JUN-1999; 99US-0333159.

(MILL-) MILLENNIUM PHARM INC.

McCarthy SA, Fraser CC, Sharp JD, Barnes TM;  
WPI; 2001-032313/04.

TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for  
screening assays and diagnostic assays and for the treatment of  
neurological diseases such as Alzheimer's, Parkinson's and Huntington's  
disease.

Claim 8; Pages 281-287; 359pp; English.

The present invention relates to TANGO or INTERCEPT proteins and coding  
sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,  
AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding  
sequences are useful for the treatment of neurological disorders such as  
central nervous system (CNS) disorders, CNS-related disorders, focal  
brain disorders, global-diffuse cerebral disorders and other  
neurological and cerebrovascular disorders. The CNS disorders include  
Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic  
lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,  
autonomic function disorders such as hypertension and sleep disorders,  
neuropsychiatric disorders, psychoactive substance use disorders,  
anxiety, and bipolar affective disorder.

Sequence 1319 AA;

Query Match 91.3%; Score 7426; DB 22; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 FNGTDLRLVNGDPCSGTVEYKFGQGTVCDDGNTTASTVVCQKLGCPSPFAMFRF 100  
Db 1 FNGTDLRLVNGDPCSGTVEYKFGQGTVCDDGNTTASTVVCQKLGCPSPFAMFRF 60

QY 101 GOAVTRHGKIWLDDVSCYGNESALWECQREHGWGSHNCYHGEDVGVNVCYGEANLGLRLVDG 160  
Db 61 GOAVTRHGKIWLDDVSCYGNESALWECQREHGWGSHNCYHGEDVGVNVCYGEANLGLRLVDG 120

QY 161 NNSCSGRVEVKFOERWGTICDDGNNLNTAAVVCROLCGPSPSFSSGVVNSPAVLRPIWLD 220  
Db 121 NNSCSGRVEVKFOERWGTICDDGNNLNTAAVVCROLCGPSPSFSSGVVNSPAVLRPIWLD 180

QY 221 DILCOGNELALWNCRRHGWGSHNCYHGEDVGVNVCYGEANLGLRLVDG 280  
Db 181 DILCOGNELALWNCRRHGWGSHNCYHGEDVGVNVCYGEANLGLRLVDG 240

QY 281 GRNGTVCCHHKNNAADVVCQKLGCGTALHFAGLPHLQSGSDVWLDGVSCSNGESFLWD 340  
Db 241 GRNGTVCCHHKNNAADVVCQKLGCGTALHFAGLPHLQSGSDVWLDGVSCSNGESFLWD 300

QY 341 CRHSGTVNFDCLHNDVSVICSDGADLELRADGNNCSGRVEVRIHEQWWTICDQNWKN 400  
Db 1319 CRHSGTVNFDCLHNDVSVICSDGADLELRADGNNCSGRVEVRIHEQWWTICDQNWKN 360

RESULT 4  
AA000396  
ID AA000396 standard; Protein; 1120 AA.  
XX  
AC AA000396;  
XX

DT 04-JUL-2001 (first entry)  
XX Human secreted protein, POLY8.  
XX Human secreted protein; therapeutic; diagnostic; human; cancer.  
XX Homo sapiens.  
XX WO2000119856-A2.  
XX 22-MAR-2001.  
XX 13-SEP-2000; 2000WO-US25106.  
XX 13-SEP-1999; 99US-0153629.  
XX 16-SEP-1999; 99US-0154520.  
XX 20-SEP-1999; 99US-0154762.  
XX 13-OCT-1999; 99US-0159231.  
XX 12-SEP-2000; 2000US-0659634.  
XX (CURA-) CURAGEN CORP.  
XX Shimkets RA, Fernandes E, Herrmann JL, Liu X, Yang M, Boldog FL;  
PI WPI; 2001-244781/25.  
XX N-PSDB; AAS01217.  
XX New POLYX polypeptide useful for treating or preventing a POLYX  
PT associated disorder, e.g. cancer -  
XX Claim 9; Page 25-29; 152pp; English.  
XX The sequence represents the amino acid sequence of human secreted  
CC protein, POLY8. POLYX nucleic acids, polypeptides and antibodies to  
CC POLYX can be used for treating or preventing a POLYX associated disorder  
CC in a subject, preferably a human. These can be used in the manufacture of  
CC a medicament for treating a syndrome associated with a human disease  
CC selected from a POLYX-associated disorder, where the therapeutic is a  
CC POLYX polypeptide, a POLYX nucleotide or a POLYX antibody. They may also  
CC be used to screen for a modulator of activity, or latency, or  
CC predisposition to a POLYX associated disorder, e.g. cancer.  
XX  
SQ Sequence 1120 AA;  
Query Match 74.4%; Score 6058; DB 22; Length 1120;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1088; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 363 DGADLELRADGNNCGRVEYRIHEQWMTICDQNNKNEQALVVCQKQCPFSVFGSRR 422  
DB 30 DGADLELRADGNNCGRVEYRIHEQWMTICDQNNKNEQALVVCQKQCPFSVFGSRR 89  
QY 423 KPSNARDIWNISICTGNESALWDCYDGGKAKRTCFRRSDAGVICSOKADLDLRLVGAH 482  
DB 90 KPSNARDIWNISICTGNESALWDCYDGGKAKRTCFRRSDAGVICSOKADLDLRLVGAH 149  
QY 483 SPYCRLEVKYOGEGTGTCHDRWTRNAAVVCQKQCPKPMHVFMTYFKEASGPIWLD 542  
DB 150 SPYCRLEVKYOGEGTGTCHDRWTRNAAVVCQKQCPKPMHVFMTYFKEASGPIWLD 209  
QY 543 VSCIGNESNIWDCEHSGKGNCHVREDIVITCSGDATWGLRLVGGSNRCSGRLEYIFOG 602  
DB 210 VSCIGNESNIWDCEHSGKGNCHVREDIVITCSGDATWGLRLVGGSNRCSGRLEYIFOG 269  
QY 603 RGTGTVDDGWNKAAAVVCQKQCPSSIGMGLGNASTGYGKIWLDDVSCDGEDSLWSC 662  
DB 270 RGTGTVDDGWNKAAAVVCQKQCPSSIGMGLGNASTGYGKIWLDDVSCDGEDSLWSC 329  
QY 663 RNSGNGNDCSHSDVGVTCSDASPMELRVGGSSRCAGKVEVNVQAGVILCANGWGN 722  
DB 330 RNSGNGNDCSHSDVGVTCSDASPMELRVGGSSRCAGKVEVNVQAGVILCANGWGN 389  
QY 723 TAEVVCRLQECGSAIRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWENKQTACHLN 782

DB 390 TAEVVCRLQECGSAIRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWENKQTACHLN 449  
QY 783 MEASLICSAPROPLRGADMPGSRVEVKHATDTRWSVCDSDFLSHAAVLCRELNGDAI 842  
DB 450 MEASLICSAPROPLRGADMPGSRVEVKHATDTRWSVCDSDFLSHAAVLCRELNGDAI 509  
QY 843 SLVSGDHFCKGNGLTWAERFQCEGSETHLALCPIVQHPEDTCTIHSREVGVVCSRYTDVRL 902  
DB 510 SLVSGDHFCKGNGLTWAERFQCEGSETHLALCPIVQHPEDTCTIHSREVGVVCSRYTDVRL 569  
QY 903 VNGKSQCQGVQVEINVLGHWSLCLDTHWDPEDARVLCRLQSCGTALSTTGKYGIGERSVRV 962  
DB 570 VNGKSQCQGVQVEINVLGHWSLCLDTHWDPEDARVLCRLQSCGTALSTTGKYGIGERSVRV 629  
QY 963 WGHREHCLGNESLLDNCOMTVLGAPCIHGNTVSVICTSLQTPLEPCLANVSDPYLSAV 1022  
DB 630 WGHREHCLGNESLLDNCOMTVLGAPCIHGNTVSVICTSLQTPLEPCLANVSDPYLSAV 689  
QY 1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFWGTICDDGWDLSDAHVVCKLGGVA 1082  
DB 690 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFWGTICDDGWDLSDAHVVCKLGGVA 749  
QY 1083 FNATVSAHFEGSGPIWLDDNLCTGTGESHLMQCPSRGWGQHDCRHKEDAGVICSEFTALR 1142  
DB 750 FNATVSAHFEGSGPIWLDDNLCTGTGESHLMQCPSRGWGQHDCRHKEDAGVICSEFTALR 809  
QY 1143 LYSETETESCAGRLEVFYNGTWGSGVRRNITTAIAGIVCRQLCGGNGVVSAPLSKTGS 1202  
DB 810 LYSETETESCAGRLEVFYNGTWGSGVRRNITTAIAGIVCRQLCGGNGVVSAPLSKTGS 869  
QY 1203 GFWMVDDIOCPKTHISIMWCLSPAPERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 1262  
DB 870 GFWMVDDIOCPKTHISIMWCLSPAPERRISSPAEETWITCEDRIRVRGGDTECSGRVEIW 929  
QY 1263 HAGSMGTVCDDSDWDLAEAEVVCQKQCGSALALRADASFGQGTGTIWLDDMRCKGNESFL 1322  
DB 930 HAGSMGTVCDDSDWDLAEAEVVCQKQCGSALALRADASFGQGTGTIWLDDMRCKGNESFL 989  
QY 1323 WDCHAKPMQSGDCGHKEDAGVRCGOSLKSALNASSGHLALILSSIFGLLLVLFILFTW 1382  
DB 990 WDCHAKPMQSGDCGHKEDAGVRCGOSLKSALNASSGHLALILSSIFGLLLVLFILFTW 1049  
QY 1383 CRVQKOKHPLRVSTRRRGSLLENLFHEMETCLKREDPHGTTRTSDTTPNHGCCDASDTSL 1442  
DB 1050 CRVQKOKHPLRVSTRRRGSLLENLFHEMETCLKREDPHGTTRTSDTTPNHGCCDASDTSL 1109  
QY 1443 LGVLPASEATK 1453  
DB 1110 LGVLPASEATK 1120  
RESULT 5  
AAU27709  
ID AAU27709 standard; Protein; 821 AA.  
XX  
XX AAU27709;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human full-length polypeptide sequence #34.  
XX  
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;  
KW nervous system disorder; inflammatory disorder; cell differentiation;  
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;  
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
KW cytostatic; antirheumatic; antiarthritic; vulnery; antinflammatory;  
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;  
KW neuroprotective; osteopathic; antidiabetic; antiallergic;  
KW immunostimulant; analgesic; gene therapy.  
XX

OS Homo sapiens.  
 PN WO200164834-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US04926.  
 XX  
 PR 28-FEB-2000; 2000US-0515126.  
 PR 18-MAY-2000; 2000US-0577409.  
 PR 17-JUN-2000; 2000US-0597707.  
 PR 14-JUL-2000; 2000US-0616807.  
 PR 19-SEP-2000; 2000US-0664641.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
 PI Drmanac R;  
 XX  
 DR WPI; 2001-589862/66.  
 DR N-PSDB; AAS44609.  
 XX  
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
 PT prepared from various human tissues, for diagnosis, treatment of  
 PT cancer, neurological, inflammatory disorders and for use in arrays for  
 PT detection  
 XX  
 PS Claim 10; SEQ ID No 206; 153pp; English.  
 XX  
 CC Sequences AAU27676-AAU28019 represent full-length polypeptides and  
 CC contig polypeptides of the invention. The proteins and their associated  
 CC DNA sequences are useful for the treatment, diagnosis and prevention of  
 CC various types of disorder in a mammalian subject such as a human, dog,  
 CC monkey, mouse, hamster or rat. The disorders include cancers such as  
 CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as  
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,  
 CC cell proliferation, cell differentiation, stem cell growth factor,  
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells  
 CC in culture to give rise to neuroepithelial cells that can be used to  
 CC augment or replace cells damaged by illness, accidental damage or genetic  
 CC disorders. The sequences may also be used for regeneration of bone  
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.  
 CC Note: Some sequences for this patent did not form part of the printed  
 CC specification, but were obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 821 AA;  
 Query Match 55.8%; Score 4540; DB 22; Length 821;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 820; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 633 MGLGNASTGKTKWLDVSCDGPESDLWSCRNSGWNNDCHSHEDVGVICSDASDMELRL 692  
 DB 1 MGLGNASTGKTKWLDVSCDGPESDLWSCRNSGWNNDCHSHEDVGVICSDASDMELRL 60  
 QY 693 VGGSSRCAGRVENVQAVGILCANGWMNIAEVVCRQLECGSAIRVSRPFHFTERTLHI 752  
 DB 61 VGGSSRCAGRVENVQAVGILCANGWMNIAEVVCRQLECGSAIRVSRPFHFTERTLHI 120  
 QY 753 LMSNSGCTGGEASLWDCIRWEWQKACHLNNEASLCSAHRQRLVGMADMPGSGRVEVKH 812  
 DB 121 LMSNSGCTGGEASLWDCIRWEWQKACHLNNEASLCSAHRQRLVGMADMPGSGRVEVKH 180  
 QY 813 ADTWRSVCDSDFSLHAANVLCRELNCDAISLSVGDHFGKNGILTWAEKFCQEGSETHLA 872  
 DB 181 ADTWRSVCDSDFSLHAANVLCRELNCDAISLSVGDHFGKNGILTWAEKFCQEGSETHLA 240  
 QY 873 LCPIVQHPEDTCTHSREVGVVCSRYTDVRLVNGKSCQDQGVENVLGHWSLCDTHWDPE 932  
 DB 241 LCPIVQHPEDTCTHSREVGVVCSRYTDVRLVNGKSCQDQGVENVLGHWSLCDTHWDPE 300  
 QY 933 DARVLCRQLSCGTALSTTGKYGIGERSVRVWGHRRFICLGNESLLDNCQMTVLGAPPCIHG 992  
 DB 301 DARVLCRQLSCGTALSTTGKYGIGERSVRVWGHRRFICLGNESLLDNCQMTVLGAPPCIHG 360  
 QY 993 NTVSVICTGSLTQPLPCLANVSDPYLSAVPEGSALICLEDKRLRLVGDGSRGAGRVEIY 1052  
 DB 361 NTVSVICTGSLTQPLPCLANVSDPYLSAVPEGSALICLEDKRLRLVGDGSRGAGRVEIY 420  
 QY 1053 HDGFWCTICDDGWDLSDAHVVCOKLGGVAFNATVSAHFEGEGSGPIWLDLNLCTGTESH 1112  
 DB 421 HDGFWCTICDDGWDLSDAHVVCOKLGGVAFNATVSAHFEGEGSGPIWLDLNLCTGTESH 480  
 QY 1113 WQCPSPRGWQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSGVRRI 1172  
 DB 481 WQCPSPRGWQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVFYNGTWGSGVRRI 540  
 QY 1173 TTAIAGIVCRQLCGGNGVVS LAPLSKTS GSGFMWDDIOCPKTHISIWOCLSA PWERRIS 1232  
 DB 541 TTAIAGIVCRQLCGGNGVVS LAPLSKTS GSGFMWDDIOCPKTHISIWOCLSA PWERRIS 600  
 QY 1233 SPAEETWITCEDIRVRGDTGTCGRVEIWHAGSWGTVCDSDWDLAEAEVVCQQLGCGSA 1292  
 DB 601 SPAEETWITCEDIRVRGDTGTCGRVEIWHAGSWGTVCDSDWDLAEAEVVCQQLGCGSA 660  
 QY 1293 LAALRASFGQGTITWLDMMRCKGNESFLWDC HAKPWGSDCGCHKEDAGVRCGSGSLKS 1352  
 DB 661 LAALRASFGQGTITWLDMMRCKGNESFLWDC HAKPWGSDCGCHKEDAGVRCGSGSLKS 720  
 QY 1353 LNASSGHLALILSSIFGLLLVLFILFTWCRVQKQKHLPLRVSTRRGSLEENLFHME 1412  
 DB 721 LNASSGHLALILSSIFGLLLVLFILFTWCRVQKQKHLPLRVSTRRGSLEENLFHME 780  
 QY 1413 TCLKREDPHGTRTSDPTPNHGCEDASDTSLGLVLPASEATK 1453  
 DB 781 TCLKREDPHGTRTSDPTPNHGCEDASDTSLGLVLPASEATK 821  
 RESULT 6  
 AAU97585  
 ID AAU97585 standard; Protein; 1151 AA.  
 XX  
 AC AAU97585;  
 XX  
 DT 27-AUG-2002 (first entry)  
 XX  
 DE Human CD163 receptor extracellular variant protein sequence.  
 XX  
 KW Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;  
 KW CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;  
 KW human immunodeficiency virus; haemolysis; haematological condition;  
 KW aplastic anaemia; iron-deficiency anaemia; megaloblastic anaemia;  
 KW sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;  
 KW lymphoma; leucopenia; splenectomy; inflammation; infection; cancer;  
 KW autoimmunity; immunodeficiency; CD163 receptor extracellular variant.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200232941-A2.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 12-OCT-2001; 2001WO-DK00671.  
 XX  
 PR 16-OCT-2000; 2000DK-0001543.  
 PR 11-JAN-2001; 2001DK-0000039.  
 PR 22-FEB-2001; 2001US-270120P.  
 XX



XX 26-DEC-2000; 2000WO-US34263.  
 XX 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HYSE-) HYSEQ INC.  
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 DR N-PSDB; AA158649.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 PT  
 XX Example 4; SEQ ID NO 2638; 10078pp; English.  
 XX  
 XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AA38642-AA42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localized neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 XX Sequence 1121 AA;  
 SQ  
 Query Match 42.1%; Score 3429; DB 22; Length 1121;  
 Best Local Similarity 56.2%; Pred. No. 2.1e-238;  
 Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;  
 QY 364 GADLELRADGNNCGRVEVRIHQWTTICDQNKNEQALVVCQKQCPFPVFGSRRAK 423  
 DB 46 GTDKELRLVDGENKCSGRVEVKVQEEWGTVCNNGWSMEAVSVCNQLGCPPTAIAKPGWAN 105  
 QY 424 PSNEARDIWINISCTGNESALWDCTYDCAKRT-CFRRSADAGVCSDKADLDLRLVGAH 482  
 DB 106 SAGSGRIWMDHVSRCGNESALWDCKHDGWHKSNCTHOODAGVTCSDGNSLEMLRTGG 165  
 QY 483 SPCYGRLEVYQGEWGTCHDRWSTRNAVAVCKQLGCGKPMHVFMTYFKEASGPIWDD 542  
 DB 166 NMCSGRIEIKFGWGTGTCDDNFNDHVASVTCQLECGSAVSFGSSNFGEGSGPIWDD 225  
 QY 543 VSCIGNESINWCEHSGWCKHNCVHREIVVTCGSDATWGLRLVGSNRCSCRLVYFQG 602  
 DB 226 LICNGESALWCKHOGWKNCHDAEDAGVTCRSGADLSRLVDPVTECSGRLEVRFG 285  
 QY 603 RWGTGDDGNSKAAAVVCSQDLPSSIGMGLGNASTGYGKIWLDDVSCDDGDESDLWSC 662  
 DB 286 ENGTICDDGWDSDYDAVACKQLGCTAVTAIGRVNASKGFGHILWDSVSCQGHFAVWQC 345  
 QY 663 RNSGWNNDCHSDVGVYTCSDADSMELRVGSSRCAGKVEVNVGAVGILCANGWGN 722  
 DB 346 KHEWGHKHCNHNEDAGVTCSDGSLRLRGSGRCAGTVEIQRLLGKVCDRGWLK 405

QY 723 IAEVVCVQLECGSAIRVSRPHFTERTLHILMSNSCGTGGASLWDCIRWEKQTACHLN 782  
 DB 406 EADVVCVQLGCGSALKTSYQVYSIQATNTWLFLLSSCNGNETSLWDCKMOWGGTCDHY 465  
 QY 783 MEASLICSARQPRLVGADMPGCSGRVEVKHADTWRSVCDSDSFLHAANVLCRELNCGDAL 842  
 DB 466 EBAKITSARHREPRLVGGDIFCSGRVEVKHGTWGSICDSDFSLEAASVLCRELQGVV 525  
 QY 843 SLSVGDHFGKGNGLTAEKFCQEGSETHALCPVIOHPEDTCIHRSREVGWVCSRYTDLVL 902  
 DB 526 SILGGAHFGEGNGQIWAEEFOCEGHEHSLCLCPVAPRPEGTCSHSRDSRVGVCSRYTEIRL 585  
 QY 903 VNGKSQCDGQVEINVLGHWGSLCDTHDPEDARVLCRQLSCGTALSTGGKVGERSVRV 962  
 DB 586 VNGKTPCEGRVELKTLGAWGLSCLNSHWDIEDAHVLCQQLKCGVALSTPGGARFGNGQI 645  
 QY 963 WGRHFCILGNESLNDNCOMTVLGAAPPCTHGNVTSVICTGSLTQPLFPCLANVSDPYLSAV 1022  
 DB 646 WRHMEHCTGTQOHMGDCPVTALGASLCTSEQVAVSICSGNSQSLSSCNSSSLGTRPTI 705  
 QY 1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVCQKLGCGVA 1082  
 DB 706 PEESAVACIESQQLRLVNGGRCAGRVEIYHEGWSGTICDDSDWLSDAHVVCRQLGCGEA 765  
 QY 1083 FNATVSAHFEGSGPIWLLDNLCTGTEHLMWQCPSPRGWQHDCHKEADGVICSEFTALR 1142  
 DB 766 INATSAHFEGGTGPIWLDLDMKCNKESRIWQCHSHGWQCNCRHKEADGVICSEFMSLR 825  
 QY 1143 LYSETETESACGRLEVFYNGTWGVSVRNITTAIAGIVCRQLGCGENGWVSLAPLSKTGS 1202  
 DB 826 LTSEASRACAGRLEVFYNGAWGTGKSSMSETTYGVVCRQLGCGADKGINPASLDKMS 885  
 QY 1203 GFMVDDIQCCKTHSIWQCLSGAPWERRISSPAETWITCEDIRIVRGSDTSCSGRVEIW 1262  
 DB 886 IPMWVDNVQCPKGPDTLWQCPSPWEKRLASPEETWITCDNKIRLQEGPTSCSGRVEIW 945  
 QY 1263 HAGSGTVCDDSDWDLAEAVVCOQLGCGSALAALDASFGQGTGTIWLDDMKCKGNESFL 1322  
 DB 946 HGSNGTVCDDSDWDLDDAQVVCQQLGCGCPALKAFAEFGQGTGTIWLNEVKCKGNESL 1005  
 QY 1323 WDCHAKPMGQSDCHGKEDAGVRCGQSLK-----SLASSGHLALILSSIFGLLLVLF 1376  
 DB 1006 WDCPARRWGHSECGHKEADAANCTDISVQKTPKATTTGRSSRQSSFFIAVGLGVLLAIF 1065  
 QY 1377 I--LFLTWCVRVOKHPLRVSTRRGSLLENLFHEMETCLAKED 1419  
 DB 1066 VALFELTKRRQROR---LAVSSRGENLVHQIYREMNCLNADD 1107  
 RESULT 8  
 AAM41279  
 ID AAM41279 standard; Protein; 1124 AA.  
 XX  
 AC AAM41279;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 6210.  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX Homo sapiens.  
 OS  
 XX WO200153312-A1.  
 PN  
 XX 26-JUL-2001.  
 PD  
 XX 26-DEC-2000; 2000WO-US34263.  
 PF







FT Domain 154..253  
 FT /label= Domain\_2  
 FT /note= "Specifically claimed in claim 6"  
 FT 261..360  
 FT /label= Domain\_3  
 FT /note= "Specifically claimed in claim 6"  
 FT 368..467  
 FT /label= Domain\_4  
 FT /note= "Specifically claimed in claim 6"  
 FT 473..572  
 FT /label= Domain\_5  
 FT /note= "Specifically claimed in claim 6"  
 FT 578..677  
 FT /label= Domain\_6  
 FT /note= "Specifically claimed in claim 6"  
 FT 714..814  
 FT /label= Domain\_7  
 FT /note= "Specifically claimed in claim 6"  
 FT 819..920  
 FT /label= Domain\_8  
 FT /note= "Specifically claimed in claim 6"  
 FT 924..1023  
 FT /label= Domain\_9  
 FT /note= "Specifically claimed in claim 6"  
 FT  
 PN WO200232941-A2.  
 XX  
 XX 25-APR-2002.  
 XX  
 PF 12-OCT-2001; 2001WO-DK00671.  
 XX  
 XX 16-OCT-2000; 2000DK-0001543.  
 PR 11-JAN-2001; 2001DK-0000039.  
 PR 22-FEB-2001; 2001US-270120P.  
 XX  
 XX (PROT-) PROTEOPHARMA APS.  
 PA  
 XX  
 XX Moestrup S, Moller HJ;  
 PI  
 XX WPI; 2002-452380/48.  
 DR  
 XX  
 PT New haptoglobin-haemoglobin complex or its mimics or receptors, useful  
 PT in gene therapy, particularly for treating an individual suffering from  
 PT e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity  
 PT  
 PT  
 PS Claim 1; Fig 5; 70pp; English.  
 XX  
 CC The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb)  
 CC complex, or its part or mimic, being operably linked to a substance,  
 CC capable of binding a CD163 receptor and/or a CD163 receptor variant.  
 CC The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor  
 CC are useful as a medicament or in the manufacture of a medicament for  
 CC treating an individual. The medicament comprises at least one drug to be  
 CC delivered to a cell expressing a CD163 receptor or a CD163 variant. The  
 CC medicament may also comprise at least one gene to be delivered to a cell  
 CC expressing a CD163 receptor or a CD163 variant. The medicament is an  
 CC antibiotic or an anticancer drug, e.g. an anti-HIV (human  
 CC immunodeficiency virus) drug. In particular, the Hp-Hb complex or its  
 CC mimic, the CD163 variant or the CD163 receptor, or the medicament are  
 CC useful for treating haemolysis and/or other haematological conditions,  
 CC aplastic anaemia, iron-deficiency anaemia, megaloblastic anaemia, sickle-cell  
 CC anaemia, polycythaemia, malaria, leukaemia, myelodysplasia,  
 CC lymphoma, leucopenia, splenectomy, inflammation, infection, cancer,  
 CC autoimmunity or immunodeficiency. The method involves drug-delivery  
 CC treatment or gene delivery treatment of an individual. The medicament is  
 CC also useful for inhibiting (in vivo) uptake of haemoglobin in cells  
 CC presenting a CD163 receptor. The CD163 variant is also useful in  
 CC diagnosing, monitoring and/or controlling a condition in an individual.  
 CC In particular, the CD163 variant is useful for identifying macrophages in  
 CC a biological sample from an individual, for identifying at least one  
 CC Hp-Hb complex in serum and/or plasma of an individual, or for removing at  
 CC least one Hp-Hb complex in serum and/or plasma of an individual. The

CC present amino acid sequence represents the human CD163 receptor of the  
 CC invention.  
 XX  
 XX  
 SQ Sequence 1116 AA;  
 Query Match 42.1%; Score 3424; DB 23; Length 1116;  
 Best Local Similarity 56.1%; Pred. No. 4.8e-238;  
 Matches 597; Conservative 161; Mismatches 295; Indels 12; Gaps 4;  
 Qy 364 GADLELRADSGNCSGRVEYRIHEQWTTICDQWKNKEALVVCVKQCPSPVFSRRRAK 423  
 Db 41 GTDKELRLVDGENKCSGRVEYKQVEEWCTVNCNGMSMAVSVICNQLGCPPTAKPAGWAN 100  
 Qy 424 PSNEARDIWINISICTGNESALWDCTYDGKAKRT-CFRRSDAGYICSDKADLDRLVGAH 482  
 Db 101 SSAGSGRIWMDHVSCRGNESALWDCKHDGKHSNCTHQDAGVTCSDGSLNEMLRTRGG 160  
 Qy 493 SPCYGRLEVYQGEWGTVCYHNRNNAVVCVKQCGCKPHVFQMTYFKASGGIWLDD 542  
 Db 161 NMSGRIEIKFQGRWGTVCDDNFNIDHASVICRQLCEGSAVSFSGSNFGEVGGPIWEDD 220  
 Qy 543 VSCIGNESNIWDCSHSGWGHKNCVHREDVIVTCSDATWGLRLVGGSNRCGRLEVFQ 602  
 Db 221 LICGNESALWNCQHGKGNCHDAEDAGYICSGADLSRLVDGVTECSGRLEVFQ 280  
 Qy 603 RWGTVCCDGMNSKAAAVVCSQDPCSSIIIGMLGNASTGYKIKWLDVSCDGEDSLWSC 662  
 Db 281 EWGTTICDDGWSYDAAVACKQLGCPPTAVTAIGRVNASKGFGHIWLDVSCQGHAPVWC 340  
 Qy 663 RNSGWNDCSHSEDEVGYICSDASDMELRLVGGSSRCAGKVEVYOGAVGILCANGWMN 722  
 Db 341 KHEWGHKHYCNHNEIDAGVTCSDGSLRLRGGGRCAGTVEVEIQRLLGKCDRGWGLK 400  
 Qy 723 IAEVVCROLEGCSAIRVSRPHFTERTLHILMSNCGCTGGEASLWDCIRWENKQFACHLN 782  
 Db 401 EADVVCRLGCGSALKTSYQVYSKIQTNTWLFSSCNGNETSLWDCRNQWGLTCDHY 460  
 Qy 783 MEASLICSAROPRLVGADMPGCSGRVEYKHAHTWRSVCDSDSFLHAANYVLCRELNGDAI 842  
 Db 461 BEAKITCSAHRPRELVGGDIPGSGRVEYKHGDTWGSICDSDSFLFAASVLCRELQCGTV 520  
 Qy 843 SLVSGDHGKGNGLTWAKFOCEGSETHALCPYVQHPEDTCIHSRGVGVCSRYTQVRL 902  
 Db 521 SILGAHFEGGNGQIWAEEFQCEGHEHSLICPVAPRPGCTCSHRDVGVCVSRYTEIRL 580  
 Qy 903 VNGKSQCDQVEINVLGHWSLCTDHPEDARVLCRQLSCGTALSTTGKGYIGERSRV 962  
 Db 581 VNGKTPCEGRVELKTLGAWGSLCNSHWDIEDAHLVLCQQLKCGVALSTPGGARFGKNGOI 640  
 Qy 963 WGRHFHCLGNESLLDNCOMTVLGAPPCIHGNTVSVICTGSLTQPLPCLANVSDPYLSAV 1022  
 Db 641 WRHMFHCTGTQEHMGDCPVTALGASLCPSQVAVSICSNQSQTLSSCSNSSLGTPRPTI 700  
 Qy 1023 PGESALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCKQLCGVA 1082  
 Db 701 PRESAVACIESQLRLVNGGGRCAGRVEIYHESGNTGTCDDSDWLDSDAHVVCRLCGGEA 760  
 Qy 1083 FNATVSAHFEGSGPIWLDLNCCTGTESHLWQCPGRGWQHCQHKEDAGVTCSEPTAIR 1142  
 Db 761 INATSAHFEGGTPTWLDKMKCNKESRIWQCHSHGNGQNCRHKEDAGVTCSEFMSLR 820  
 Qy 1143 LYSETETESACRLEVFYNGTWGVSRRNITTAIAGVLCRQLCGGNGVVSVLAPLSKTGS 1202  
 Db 821 LPSEASREACARLEVFYNGAWGTVGKSSMETTVGVVCRQLCGCADKGINFASLDKAMS 880  
 Qy 1203 GPMVDDDIQCPHTHISIOCLSPAWERRISSPAEETWITCEDRIYRVGGDTCCSGRVEIW 1262  
 Db 881 IPWVDVNVQCPAGPTLMQCPSPWPKRLASPSSEETWITCDNKRLEQGTSCSGRVEIW 940  
 Qy 1263 HAGSGWGTVCDDSDWDLAEAEVWCQQLCGCGSALALRADASFGQGTGTWLDMDCKGNESFL 1322  
 Db 941 HGSWGTVCDDSDWDLDDAQQVVCQQLCGCPALKAFAEFQGGTGTPTWLNVEVCKGNESL 1000

QY 1323 WDCHAKPWGSDGCHKEDAGVRCGSLK-----SLNASSGHLALILSSIFGLLLVLF 1376  
 Db 1001 WDCPARWGHSECHKEDEAAVNCCTDISVQKTPQKATGRSSROSSFTAVGLIGVLLAIF 1060  
 QY 1377 I--LFLTWCRVQOKHPLPLVSTRRGSLLENLFPHENETCLKRED 1419  
 Db 1061 VALFFLTKRRQRQ---LAVSSRGLENLVHQIYREMNCLNADD 1102

RESULT 12  
 AAU97583  
 ID AAU97583 standard; Protein; 1149 AA.  
 XX  
 AC AAU97583:  
 CC  
 DT 27-AUG-2002 (first entry)  
 XX  
 DE Human CD163 receptor cytoplasmic variant 1 protein sequence.  
 XX  
 KW Human; CD163; receptor; haptoglobin-haemoglobin complex; Hp-Hb complex;  
 KW CD163 receptor variant; antibiotic; anticancer drug; anti-HIV drug;  
 KW human immunodeficiency virus; haemolysis; haematological condition;  
 KW aplastic anaemia; iron-deficiency anaemia; megaklastic anaemia;  
 KW sickle-cell anaemia; polycythaemia; malaria; leukaemia; myelodysplasia;  
 KW lymphoma; leucopenia; splenectomy; inflammation; infection; cancer;  
 KW autoimmunity; immunodeficiency; CD163 receptor cytoplasmic variant 1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W020232941-A2.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 12-OCT-2001; 2001WO-DK00671.  
 XX  
 PR 16-OCT-2000; 2000DK-0001543.  
 PR 11-JAN-2001; 2001DR-0000039.  
 PR 22-FEB-2001; 2001US-270120P.  
 XX  
 PA (PROT-) PROTEOPHARMA APS.  
 XX  
 PI Moestrup S, Moller HJ;  
 XX  
 DR WPI; 2002-452380/48.  
 XX  
 PT New haptoglobin-haemoglobin complex or its mimics or receptors, useful  
 PT in gene therapy, particularly for treating an individual suffering from  
 PT e.g. haemolysis, anaemia, inflammation, infection, cancer or autoimmunity  
 XX  
 PS Claim 1; Fig 5; 70pp; English.  
 XX  
 CC The present invention relates to a new haptoglobin-haemoglobin (Hp-Hb)  
 CC complex, or its part or mimic, being operably linked to a substance,  
 CC capable of binding a CD163 receptor and/or a CD163 receptor variant.  
 CC The Hp-Hb complex or its mimic, the CD163 variant or the CD163 receptor  
 CC are useful as a medicament or in the manufacture of a medicament for  
 CC treating an individual. The medicament comprises at least one drug to be  
 CC delivered to a cell expressing a CD163 receptor or a CD163 variant. The  
 CC medicament may also comprise at least one gene to be delivered to a cell  
 CC expressing a CD163 receptor or a CD163 variant. The medicament is an  
 CC antibiotic or an anticancer drug, e.g. an anti-HIV (human  
 CC immunodeficiency virus) drug. In particular, the Hp-Hb complex or its  
 CC mimic, the CD163 variant or the CD163 receptor, or the medicament are  
 CC useful for treating haemolysis and/or other haematological conditions,  
 CC aplastic anaemia, iron-deficiency anaemia, megaklastic anaemia,  
 CC sickle-cell anaemia, polycythaemia, malaria, leukaemia, myelodysplasia,  
 CC lymphoma, leucopenia, splenectomy, inflammation, infection, cancer,  
 CC autoimmunity or immunodeficiency. The method involves drug-delivery  
 CC treatment or gene delivery treatment of an individual. The medicament is  
 CC also useful for inhibiting (in vivo) uptake of haemoglobin in cells  
 CC presenting a CD163 receptor. The CD163 variant is also useful in  
 CC diagnosing, monitoring and/or controlling a condition in an individual.

CC In particular, the CD163 variant is useful for identifying macrophages in  
 CC a biological sample from an individual, for identifying at least one  
 CC Hp-Hb complex in serum and/or plasma of an individual, or for removing at  
 CC least one Hp-Hb complex in serum and/or plasma of an individual. The  
 CC present amino acid sequence represents the human CD163 receptor  
 CC cytoplasmic variant 1 of the invention.  
 XX  
 SQ Sequence 1149 AA;  
 Query Match 41.7%; Score 3396.5; DB 23; Length 1149;  
 Best Local Similarity 54.4%; Pred. No. 4.8e-236;  
 Matches 597; Conservative 161; Mismatches 295; Indels 45; Gaps 5;  
 QY 364 GADLERLADGSNNCSGRVEVRIHEOWMTICDQNKNEQALVVKQGCFFSVFSGSRAK 423  
 Db 41 GTDKELRLVDGENKCSGRVEVKVQEEGTVCNNGMSMEAVSVICNLQCTPAIRAPGAN 100  
 QY 424 PSNEARDIWNISICTNESALWDCTVDGKAKRT-CFRSRDAGVCSKADLRLVGAH 482  
 Db 101 SSAGSGRIWMDHVSRCGNESALWDCKHDGKMGKHSNCTHQDAGVTCSDGSLNMLRTGG 160  
 QY 483 SPYGRLEVKYQGEWGTVCWHDWSTRNAAVVCKOLGCGKPMHVGMTYFKEASGPILWDD 542  
 Db 161 NMSGRIEIRPQGRWGTVCDDNFIDHASVTCROLEGSVFSFGSSNFGGSGPIWDD 220  
 QY 543 VSCIGNESNIWDCSHSGWGHKNCVHREDVITVCSGDATWGLRLVGGNRCSGRLEVIFQG 602  
 Db 221 LICNGESALWNCQHGKHCNCDHAEDAGVICSAGDLSRLVDGVTECSGRLEVIFQG 280  
 QY 603 RWGTVCDDGWNKAAAVVCSQDPCSSIIGMLGNASTGYGKIWLDDVSCDSDSLWSC 662  
 Db 281 EMGTICDDGWDSDYDAAVACKOLGCGPTAVTAIGRVNASKGFGHILWDSVSCQGHPEAVWOC 340  
 QY 563 FNSGWGNDCHSHSEDEVGIVCSDSADMELRLLVGGSSRCAGKVEVNVQAVGILCANGWGN 722  
 Db 341 KHHEWGHYCNHNEEDAGVTCSDGDLRLRGGGSRCACTVEVEIQRLLGVCYDRGWLK 400  
 QY 723 IAEVVCROLECGSAIRVSRPHEFTERTLHILMSNSGCTGGSEASLWDCIRWEMKQATACHLN 782  
 Db 401 EADVVCROLECGSALKTSYQVYSKIQTNTWLFSSCNGNETSLWDCNKNWQGLTCDHY 460  
 QY 783 MEASLICSAPRQRLVGMADPCSGRVEVKHADTWRSVCDSDSFLHAANVLCRELNCDAI 842  
 Db 461 EEAKITCSAHPRLVGGDIPCSGRVEVKHGTWGSICDSDSFLHAANVLCRELNCDAI 520  
 QY 843 SLSVGDHFGKNGLTWAEKFOCESETHLALCPVQHPEDTCIHSREVVVCSR----- 896  
 Db 521 SILGGAHFGEGNGOIWAEFOCEGESHLSLCPVAPRPEGTCSHSRDVGVVVCSSKTOKTS 580  
 QY 897 -----YTDVRLVNGKSCDQVQVEINVLGHWSLCIDTHW 929  
 Db 581 LIGSYTVKGTGLGSHSCLFLKPCLLPGYTEIRLVNGKTPCEGRVELKTLGAWGSLCNHW 640  
 QY 930 DPEDARVLCRLSCGTALTSTGGKYIGERSVVRVNGHRPHCLGNESLNDNCMTVLGAPPC 989  
 Db 641 DIEDAHLVLCQQLKCGVALSTPGGARFGKNGQIWRHMFHCTGTGEOHMGDCPVTALSLC 700  
 QY 990 IHGNTVSVICTGSLTQPLFPCLANVSDPYLSAYPEGSALICLEDKRLRLVGDSPRCAGV 1049  
 Db 701 PSEQVAVSICSGNSQTLSSCNSSLGPTPTPEESAVACIESQGLRLVNGGRCAGV 760  
 QY 1050 EYHDGFWGTICDDGWDLSDAHVVCQKCGVAFNATVSAHFEGSGPIWDDLNCTGTE 1109  
 Db 761 EYHEGSGWTICDDGWDLSDAHVVCRLGCGEATNATGSAHFEGGTGPIWDDKCNKE 820  
 QY 1110 SHLWQCPSRGWHQDCHRHKEDAGVICSEFTALRLYSETETESACARLEVFNWGTWGSYGR 1169  
 Db 821 SRIWQCHSHGWQCNCRHKEDAGVICSEFTALRLYSETETESACARLEVFNWGTWGSYGR 880  
 QY 1170 RNITATAGIVCROLGCGGCVWSLAPLSKTCGFMWVDDIQCPTKTHISIWQCLSAWHER 1229  
 Db 881 SSMSETTVGVVVCRLGCGADKGNKIPASLDKAMSIPMWVNDVQCPKPTDLWOCPPSPWEK 940



Db 886 SSMSTTVGVVCRQLGCKADKGINPASLDKAMISPMWVDNYQCPKGPDTLWQCPSPWEK 945

Qy 1230 RISSPAETWTCEDRIRVRGGDTCSGRVEIWHAGSNGTVCCDSDSLAEAEVWCQQLGC 1289

Db 946 RLASPESETWTCDNKIRLQGGPTSCSRVEIWHGSGMGTVCDDSDLDLDAQVVCQQLGC 1005

Qy 1290 GSALAALRDASFGQGTGTIWLDDMRCKNGESFLWDCHAKPMGQSDCGHKEBAGVRCSSQS 1349

Db 1006 GPALKAFKAEAFEGGTGTIWLNEVKCKNESSLWDCPARRNGHSCGHKEBAAVNCNTDIS 1065

Qy 1350 LK-----SLNASSGHLALIUSSIFGLLLLVLF--LFTWCRVQKQKHLPLRVSTRRG 1401

Db 1066 VQKTPQKATTGRSSQSSFIAGVILGVVLLAIFVALFELTKRRQRQ--LAVSSRGEN 1122

Qy 1402. SLEENLFHEMETCLKRED 1419

Db 1123 LVHQIQYREMNLSCLNADD 1140

RESULT 14

AA666088

ID AA666088 standard; Protein: 1436 AA.

XX AA666088;

AC AA666088;

DT 30-MAR-2001 (first entry)

XX Bovine WCI protein.

DE

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder; central nervous system; focal brain disorder; bipolar affective disorder; global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease; senile-dementia; Huntington's disease; amyotrophic lateral sclerosis; Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep; neuropsychiatric; psychoactive substance use; anxiety.

XX Bos sp.

OS

XX WO200077239-A2.

XX

XX 21-DEC-2000.

PD

XX 24-MAY-2000; 2000WO-US14858.

XX

XX 14-JUN-1999; 99US-0333159.

XX

XX (MILL-) MILLENNIUM PHARM INC.

XX

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

PI

XX WPI; 2001-032313/04.

XX

XX N-PSDB; AAF45141.

DR

XX

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for screening assays and diagnostic assays and for the treatment of neurological diseases such as Alzheimer's, Parkinson's and Huntington's disease -

PT

XX

XX Disclosure; Fig 2; 359pp; English.

PS

XX

XX The present invention relates to TANGO or INTERCEPT proteins and coding sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057, AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding sequences are useful for the treatment of neurological disorders such as central nervous system (CNS) disorders, CNS-related disorders, focal brain disorders, global-diffuse cerebral disorders and other neurological and cerebrovascular disorders. The CNS disorders include Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome, autonomic function disorders such as hypertension and sleep disorders, neuropsychiatric disorders, psychoactive substance use disorders, anxiety, and bipolar affective disorder. The present sequence is a sequence used in a sequence homology alignment with the TANGO/INTERCEPT

CC sequences of the present invention.

XX

SO Sequence 1436 AA;

Query Match 37.9%; Score 3082; DB 22; Length 1436;

Best Local Similarity 41.7%; Pred. No. 3.le-213;

Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;

Qy 28 CILLINSCFLSSFNPTDLELRVLVNGDPCGCTVEVKFQGGMTVCDDGWNNTASTVVK 87

Db 13 CVLLIG-----TWGGQALELRKLDGVHRCEGRVEVKKHGGMTVDGRTWTLKDSVYVCR 67

Qy 88 QLGCPFSFAM---FRFGQAVTRHGKIWLDDVSCYNESALMECQH---REMGSHNYHGE 141

Db 68 QLGCAAGIFPGGAYFGPL---GPIWLLYTSCETESTVSDCEHSNIDKYRNDYNGHR 124

Qy 142 DGVNVCYGEANLGLRLVDGNNSCGRVEVKFOERMGTTCCDGGNUNLNTAAVYCRQLGCPSS 201

Db 125 DAGVVCSG---FVRLAGGDGPCSGRVEVHSGEAMIPVSDGNFTLATQAIIICAEELGCGKA 180

Qy 202 FISSGVVNSPAVLRPIWLDDILCOGNELALNCRHGRGNHDCSHNEDVTLCYDSSDLE 261

Db 181 VSVLGHLEFPRESSAQVWAEFRCEGEPELWVCPVPCPGGTCCHHSGSAQVVCSEYR 240

Qy 262 LRLVGTNRCMGRVELKIQGRMGTVCHHKWNNAAADVVCQKOLGCGTALHFAGLPHLQSGS 321

Db 241 L-MTNGSSQCEGQVEMNISGQWRALCASHWSLANANVICRLGCGVAISTPGPHLVEEG 299

Qy 322 DVWLDGVSCSGNESFLWDCRHSHTVNFDCIHLQNDVSVICSDGADLELRDLADGNSNCSR 381

Db 300 DQILTARFHCSGAESFLMSCPVTALGDPDCSHGNTASVICSGN----- 342

Qy 382 YEVRIHEQWMTICDQNMKNQALVVCOLGCPFSVFGSRRAKPSNEARDIWNISICTGN 441

Db 343 -QIQVLPO-----CNDSV-----SQTSAAS----- 363

Qy 442 ESALWDCITYDGKAKRTPCFRRSDAGVICSADKDLRLVLAHSPCYGRLEVKYQGEWTV 501

Db 364 -----EDSAPYCSDSR--QLRLVDGGGPGCAGRVEILDQSGWGTIC 401

Qy 502 HDRWSTRNAAVVVCQKLGCGKPMHVFMTYFKEASGPILWLDVSCIGNESNIWDCSHSGW 561

Db 402 DDGNDLDDARVVCQKLGCGEALNATGSAHFGAGSGPIWLDNLNCTGKESHVWRCPSRGW 461

Qy 562 KNCVHREDVITVTCSDATWGLRLVGGSNRCSGRLEVYFGRWGTVCDDGWNKAAAVVC 621

Db 462 QNCRHKQDAGVICS--EFLALRWVSEDOQACAGLEVYNGTWGVCNRPMDIIVSTIC 519

Qy 622 SQLDCPSSIIGMLGNAST---GYGKIWLDDVSCDGEDSLWSCRNSGWNDCSHSDE 677

Db 520 ROLGCGDS---GTLSNVALLREGFRPWVDRIQCRKTDTSWQCPSPDPWYNNSCPKEE 575

Qy 678 VGVICSDASDMLRLVGGSSRCAGVENVQGAIVILCANGHGMNIAEVVCQRLGCGSAI 737

Db 576 AYIWCADS--RQIRLVDDGGRCGRVEILDQGSWGTICDDRWDLDDARVVCQKLGCGEAL 633

Qy 738 RVSREPHPTERTLHILMSNGCTGCEASLWDCIRWEKQTCACHLNEASLICASAHQRPL 797

Db 634 DATVSSFFGTGSGPTLWDEVNCRGEESQVWRCPSPGWGRHNCNHQEDAGVICSGF--VRL 691

Qy 798 VGADMPGCSGRVEVKHADTWRSVCDSDFLHAAANVLCRELNCGDAISLSVGDHFGKNGUT 857

Db 692 AGGDGPGCSGRVEVHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSVLGHMPFRESQGV 751

Qy 858 WAEKFCQCEGSETHLALCPVQHPEDTCIHSREVGVVCSRYTDVRLV--NGKSCDQVLEIN 916

Db 752 WAEERFCDDGGEPELWSCPVRPCPGTCLHSGAAVVCVSYTTEVQLMKNGTSQCEQVENK 811

Qy 917 VLGHWSLCLDTHWDPEDARVLCRLQSCGTALSTTGKVIIGERSVRVWGRHRCILGNESLL 976

Db 812 LSGRRALCASHWSLANANVVCRLGCGVAISTPGPHLVEGGDQISTAQFHCSGAESFL 871

Qy 977 DNQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRL 1036

Db 872 WSCPVTALGGPCHSHGNTASVTCGNHTQVLPQCNDFLSQPAGSAASESSPYCSDSROL 931  
 QY 1037 RLVDGSRACAGVEYTHDGFMTICDDGDWDLSDAHVVCOKLGGVAFNATVSAHFGEGSG 1096  
 Db 932 RLVDGGCGGVRVEILDQGSWTICDDWDLDDARVVCROLGGGALNATGSAHFGAGSG 991  
 QY 1097 PIWLDLNLCTGTFESHLMQCPGSRGWQHCRCRKEDAGVICSEFTALRLYSETESETSCAGRL 1156  
 Db 992 PIWLDLNLCTGKESHVWRCPGSRGWRHDCRKEDAGVICSEFTALRLMVS--EDQQCAGWL 1049  
 QY 1157 EYFYNCTWGSVGRNRTTAAIGVRCQLCGGNGVVS LAPLSKTCGSGFMVDDIOCPKTH 1216  
 Db 1050 EYFYNCTWGSVGRSPMEDITVSVICRLCGDGSGLNTSVGLREGSRPRWDLIOCRKMD 1109  
 QY 1217 ISTWQCLSAPEWRRISSPAETWITCE-----DRIRVRGGTECSGRVE 1260  
 Db 1110 TSLWQCPSPGFWKYSSCPKEAYISCEGRRPKSCPTAAACTDREKLRLGGSECSGRVE 1169  
 QY 1261 INHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALALRDASFQGGTGTIWLDDMRCKNGNS 1320  
 Db 1170 VHWGSMGTVCDDSWSLAEAEVVCQOLGCGQALEAVRSAAFPGNGSIWLDEVCQGGRES 1229  
 QY 1321 FLWDCHAKPWGQSDCGHKEDAGVRCSG-----QSLKSLNASGHLALIL 1364  
 Db 1230 SLWDCVAEPWGQSDCKHEEDAGVRCGVRITLPTTTAGTRTTSNLPGLFSLPGVLCIL 1289  
 QY 1365 SSIFGLLLVFLFTLCWRCVOKHPLRVSTRRRRGSLEENLFHEMETCLKREDPHGTR 1424  
 Db 1290 GSLFLVLVILYQLRW-RAER-----RALSSYEDALAEAYEBLDYLLTQKEGLGSP 1342  
 QY 1425 TS-DTTPNRCEDASDTSLGLVLPASEATK 1453  
 Db 1343 DQMTDVPDENYDAAEVPVPGTPSPSQGNE 1372

## RESULT 15

ID AAW64591  
 AC AAW64591 standard; Protein; 1785 AA.

XX AAW64591;  
 XX 23-OCT-1998 (first entry)

DT Human SRCR protein.

DE Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;  
 KW nervous system; medullo-blastoma; glioma; breast; detection;  
 KW antibody; ss.

XX Homo sapiens.

XX WO9830687-A2.

PD 16-JUL-1998.

XX 09-JAN-1998; 98WO-DE00096.

XX 18-JUL-1997; 97DE-1030997.

PR 09-JAN-1997; 97DE-1000519.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Mollenhauer J, Poustka A;

PI WPI: 1998-399136/34.

DR N-PSDB; AAV49652.

XX Proteins containing scavenger receptor, cysteine rich domain -

PT useful for diagnosis and treatment of tumours

PS Claim 2; Fig 2; 54pp; German.

XX

CC This sequence represents a human protein which contains a SRCR (scavenger  
 CC receptor, cysteine-rich) domain. The gene and encoded protein can be used  
 CC to diagnose or treat tumours, particularly of the nervous system  
 CC (medullo-blastoma or glioma) or breast. The DNA sequence and probes  
 CC derived from it, are used to identify genes that express SRCR-domain  
 CC containing proteins, to determine the form in which these proteins exist  
 CC and to assess the significance of individual forms on cellular  
 CC properties. The protein can be used to detect the presence of  
 CC autoantibodies and antibodies which regulate its expression.

XX Sequence 1785 AA;

SQ Query Match 29.9%; Score 2435; DB 19; Length 1785;  
 Best Local Similarity 36.1%; Pred. No. 1.9e-166;  
 Matches 533; Conservative 155; Mismatches 415; Indels 374; Gaps 28;

QY 43 GTD--LELRVNGDCPGSGTVEVKFGQWGTVCDDGWNNTASTVVCVKQLGCPFSFAM--- 97  
 Db 95 GDSGLALRLVNGDRCQGRVEILYRGSGTVCDDSWDTNDANVVCRLGCGWAMSAPON 154  
 QY 98 FRFGQAVTRHGKIWLDDVSCYGNESALMPCQHREWGSHNCHYHGEDVGVNCHY 148  
 Db 155 ANFGQG---SGPIALDDVRCSGHESYLWSCPHNGWLSHNCGHGEDAGVICSAQAPOSTLR 211  
 QY 149 -----GEANLGLRLVDGNNCSGRVEYKFOERWGTICDDGWNLTAAVY 192  
 Db 212 PPSVPRISPPVPTGSESSALRLVNGDRCGRVEILYRGSGTVCDDYWDNDANVY 271  
 QY 193 CRLGCPSPSFISSGVNNSPAVLRPIWLDDILCQGNELALWNCRHNGNHDCHSHNEDVTL 252  
 Db 272 CRLGCGWAMSAPGNAQFGQSGPIVLDDVRCSGHESYLWSCPHNGWLTCHNCGHSEDAGV 331  
 QY 253 TCY-----DSSDLRLRLVGGTNRCMGRVELKTOGRWGTVCCHKH 290  
 Db 332 ICSPQSRPTSPDPTWPTSHASTAGPESSALRLVNGDRCGRVEILYRGSGTVCDDSD 391  
 QY 291 WNNAAADVVCKLGGCTALHFAGLPHLQSGSDVWVLDGVSCSGNESFLWDCRHSGTVNED 350  
 Db 392 WDTSDANVVCRLGCGWATSAFNGARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHN 451  
 QY 351 CLHONDVSVICSDGADLELRADGNNCSGRVEIRHEQWWTICQGNWNEQALVVCCKQL 410  
 Db 452 CQHSEDAGVICSA-----HSWST----- 470  
 QY 411 GCFPSVFGSRRAKPSNEARDIWNISCTGNESALWDCITYDGKAKRTCFRRSDAGVICSD 470  
 Db 471 -----PSPDTPITLTPASTVGSSESS----- 491  
 QY 471 KADLDRLVCAHSPCYGRLEVYKQGEWGTVCCHDRWTRNAAVVCVKOLGCGKPMHVFGMTY 530  
 Db 492 ---LALRLVNGDRCQGRVEILYQGSWGTVCDDSDWDTNDANVVCRLGCGWAMSAPGNAR 548  
 QY 531 FKEASGPIWLDDVSCICIGNESNIWDCSHSGKGNCHNCVHREDVIVTCSDGA-----TW- 581  
 Db 549 FGQSGPIVLDDVRCSGHESYPWSCPHNGWLSHNCGHSEDAGVICSAQSRPTSPDTPW 608  
 QY 582 -----GLRLVGSNRCSGRLVEYFQGRWGTVCDDGWNNSKAAAVVCSQLDCPS 628  
 Db 609 TSHASTAGSESSALRLVNGDRCQGRVEILYRGSGTVCDDYWDNDANVVCRLGCGW 668  
 QY 629 SIIGMGLGNASTCYGKIWLDDVSCDGEDSLWSCRNSGNNDCHSDSEVDGVICS----- 683  
 Db 669 AMSAPGNARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICSAQSQ 728  
 QY 684 -----DASDMELRLVGGSSRCAGKVEYNVOGAVGILCANGMGMIAEV 726  
 Db 729 PTPSPDTPWPTSHASTAGSESSALRLVNGDRCQGRVEILYRGSGTVCDDYWDNDANV 788  
 QY 727 VCRQLCEGSAIRVSRPHPTERTLHILMNSGCTGGEASLWDCIRWEWQKTACHLWNEAS 786  
 Db 789 VCRQLGCGWATSAFNGARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHEDAG 848  
 QY 787 LICSAHR-QPRLVGADMPGCSGRVEYKHAOTWRSVCSDFSLHAANVLCRELNCGDAISLS 845

```
Db 849 VICASOSQP-----TPSPDTWPT----- 867
Qy 846 VGDHFGKNGLTWAEKFCQEGSETHALCPVQHPEDTCIHSREVGVVCSRYTDOVRLVNG 905
Db 868 -----SRASTAGSESTLAL-----RLVNG 886
Qy 906 KSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRLQSCGTALSTGGKIYIGERSVRVWGH 965
Db 887 GDRCKGRVEVLYQGSWGTVCDDYWDTNANVVCRLCGWAMSAPNAQFGGSGPIVLD 946
Qy 966 RFHCLGNESLLDNQMTVLGAPPCIHGNTVSVICTSLTO--PLFPCLANVSDPYLSAVP 1023
Db 947 DVRCSGHESYLWSCP HNGHSHNGHEDAGVCSAAQSQSTPRPDTWLTTLNLPALT VGS 1006
Qy 1024 EGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGDWLSDAHVVCOKLGCQVAF 1083
Db 1007 ESSL-----ALRLVNGDRCRVEVLYRGSWGTVCDDSDWTDNANVVCRLCGCWAM 1059
Qy 1084 NATVSAHFEGSGPIWLDDLNCTGTESHLMQCPGRGWGQDCHRKEDAGVICSEFTALRL 1143
Db 1060 SAPGNAREFGSGPIVLDDVRCSGNESYLMSCPHKWLTHNCGHHEDAGVICS---ATQI 1116
Qy 1144 YSET-----ETESCAGRLEVY-NGTWGS----- 1166
Db 1117 NSTTDDWHPHTTTTARPSNCGGFL--FYASGTFFSPSPYYPNNAKCVWEIEVANSY 1174
Qy 1167 ---VGRNI-----TTATAGVC---ROLGCGENGCVSL---AP 1196
Db 1175 RINLGFNLKLEAHNCSFDYEIFDGSLSNLSLLGKICNDTROIFTSSYNRMTHFRSD 1234
Qy 1197 LSKTSGGFM-WVDDIOCPKTHISIWQCLSAFWERRISSPAEETWITCEDRIRVRGGDTEC 1255
Db 1235 ISFQNTGFLAWN-----SFPSDATL-----RLVNLNSSYCLC 1267
Qy 1256 SGRVEIHWAGSWGTVCDSDWDLAEAVVVCQOLGCGSALALRDASFGQGTGTIWLDDMRC 1315
Db 1268 AGRVEIYHGGTWGTVCDDSWTIQAEVVCROLGCGRAVSALGNAYFGSGGPIITLDDVEC 1327
Qy 1316 KGNESFLWDCHAKPWGQSDCGHKEKEDAGVRCSGQSLKS 1352
Db 1328 SGTESTLWQCRNRCWFHNCNHRDAGVICSGNHLST 1364
```

Search completed: May 12, 2003, 06:37:17  
Job time : 63 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:29:56 ; Search time 33 Seconds  
(without alignments)  
4232.828 Million cell updates/sec

Title: US-09-759-130B-381  
Perfect score: 8138  
Sequence: 1 MMLPQNSWHDGRCCHQHN.....CEDASDTSLGLVPASEATK 1453  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: piri:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3441.5	42.3	1151	I38004	M130 antigen precu
2	3429	42.1	1156	I38005	M130 antigen precu
3	3396.5	41.7	1149	I38006	M130 antigen precu
4	3082	37.9	1436	A46496	antigen WCI.1 prec
5	2990	36.7	2403	A59386	sanko - human
6	2439.5	30.0	2153	T14893	scavenger receptor
7	2246.5	27.6	2043	T18524	scavenger receptor
8	2156.5	26.5	1594	T30549	hensin - rabbit
9	1962.5	24.1	2083	T42721	CRP-ductin-alpha p
10	1076.5	13.2	600	S56744	mucin (clone pGM7 -
11	1048.5	12.9	1290	A57190	ebnerin precursor
12	866.5	10.6	1036	T17405	scavenger receptor
13	720.5	8.9	761	JC5759	brain-specific ser
14	665	8.2	918	JC4361	scavenger receptor
15	626.5	7.7	626	I49100	mscd6 precursor -
16	594.5	7.3	468	S26741	T-cell glycoprotein
17	520	6.4	532	A32751	speract receptor p
18	385	4.7	504	S56745	mucin (clone pGM31
19	314.5	3.9	577	A53202	cyclophilin C-asso
20	311.5	3.8	585	A47161	Mac-2-binding glyc
21	290.5	3.6	538	A55840	macrophage bacteri
22	290	3.6	458	B44407	macrophage scaveng
23	288	3.5	451	A38415	macrophage scaveng
24	279.5	3.4	453	S08276	macrophage scaveng
25	279.5	3.4	454	I46862	macrophage scaveng
26	273.5	3.4	454	A44407	macrophage scaveng
27	264	3.2	494	A29079	lymphocyte surface
28	258	3.2	5376	T42215	zonadhesin - mouse
29	254.5	3.1	495	A26396	T-cell surface gly

cell-fate determin  
Notch-1 protein -  
fibrillin-1 precu  
MEGF6 protein - ra  
alpha-2-macroglobu  
notch protein homo  
T-cell surface gly  
notch homolog - se  
Balbiani ring 3 pr  
Notch homolog prot  
fibrillin 1 precu  
notch protein - fr  
fibrillin I - bovi  
alpha-2-macroglobu  
fibrillin-2 precu  
T-cell surface gly.

ALIGNMENTS

RESULT 1

I38004

M130 antigen precursor, splice form 3 - human

N:Alternate names: M130 antigen cytosolic variant 1

C:Species: Homo sapiens (man)

C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 17-Mar-2000

C:Accession: I38004; S36078

R:Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D.  
Eur. J. Immunol. 23, 2320-2325, 1993

A:Title: A new macrophage differentiation antigen which is a member of the scavenger

A:Reference number: I38003; MUID:93380506; PMID:8370408

A:Accession: I38004

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1151 <LAW>

A:Cross-references: EMBL:Z22969; NID:g312143; PIDN:CAA80542.1; PID:g312144

C:Comment: For other splice forms, see PIR:I34005 and PIR:I34006.

C:Genetics:

A:Gene: GDB:M130; GDB:MM130

A:Cross-references: GDB:119369; OMIM:133710; GDB:9958989

A:Map position: 10pter-10qter

C:Note: the locus has not been clearly established and there is some confusion in the

C:Superfamily: scavenger receptor cysteine-rich domain homology

C:Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protei

F:1-40/Domain: signal sequence #status predicted <SIG>

F:41-1151/Product: M130 antigen, splice form 3 #status predicted <MAT>

F:41-1045/Domain: extracellular #status predicted <EXT>

F:43-147/Domain: scavenger receptor cysteine-rich domain homology <SRC1>

F:151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC2>

F:258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>

F:365-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>

F:470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC5>

F:575-678/Domain: scavenger receptor cysteine-rich domain homology <SRC6>

F:711-814/Domain: scavenger receptor cysteine-rich domain homology <SRC7>

F:816-920/Domain: scavenger receptor cysteine-rich domain homology <SRC8>

F:921-1024/Domain: scavenger receptor cysteine-rich domain homology <SRC9>

F:1046-1066/Domain: transmembrane #status predicted <TRN>

F:1067-1151/Domain: intracellular #status predicted <INT>

F:100,118,226,315,440,680,689,762,996/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 42.3%; Score 3441.5; DB 2; Length 1151;

Best Local Similarity 55.2%; Pred. No. 2.1e-215;

Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

QY 364 GADLELRADGNNCSGRVEVRIHEQWMTICDQNKNEQALVQKLGCPFSVFGSRRAK 423

Db 41 GTDKELRLVDGKNCGRVEVKVQEWCTVCNNGNMEAVSVCNLCGCTAIKAPGAN 100

QY 424 PSNEARDIWNISCTGNSALWDCTVDGKAKRT-CFRSDAGVICSADKDLRLVGAAH 482

Db 101 SSAGSGRIWMDHVHSCRGNSALWDCKHDMGKHSNCTHQDQAGVTCSDGSLNLEHRLTRGG 160

Qy 483 SPYGRLEVKYQGEWGTVDHWRWTRNAAVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542  
Db 161 NMCGRLEIKFGQGWGTVDHWRWTRNAAVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 220  
Qy 543 VSCIGNESINWDCSHSGKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602  
Db 221 LTCNGESALNCKHOGKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 280  
Qy 603 RWTGTCDDGWSNKAANVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542  
Db 281 EWGTCDDGWSNKAANVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542  
Qy 663 RNSGWNDCSHSGKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602  
Db 341 KHEWGHKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602  
Qy 723 TAEEVCRQLECGSAIRVSRPHTERTLHILMSNGSGTGGASLWDCIRWENKQTAACHLN 782  
Db 401 EADVVCRLQCGSALTKSYQVYSKIQTANTWFLSSCNGNETSLWDCIRWENKQTAACHLN 782  
Qy 783 MEASLICSARHROPRLVGADMPGSRGVKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602  
Db 461 EAKITCSARHROPRLVGADMPGSRGVKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602  
Qy 843 SLVSGDHFCKGNGLTWAEKFCQEGSETHALCPVIOHPEDTCIHSREVGVCSTYDRL 902  
Db 521 SILGAGHFGEGNQIWAEEFQEGHESHLSCPVAPRPGTCSHRSRGGVVCSTYDRL 902  
Qy 903 VNGKSCDQGVINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGGYIGERSVRV 962  
Db 581 VNGKTPCEGRVELKTLGAWSLCNSHWDIEDAHVLCQQLKCGVALSTPGGARFGKNGQI 640  
Qy 963 WGRHFCNLGNSLNDNCOMVTGLPAPCTHGNVTVCISLTQPLPCLANVSDPYLSAV 1022  
Db 641 WGRHFCNLGNSLNDNCOMVTGLPAPCTHGNVTVCISLTQPLPCLANVSDPYLSAV 1022  
Qy 1023 PEGSALICLEDRLRLVGDGSRACGRVEIYHDFGWTGTCDDGWSNKAANVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542  
Db 701 PEGSALICLEDRLRLVGDGSRACGRVEIYHDFGWTGTCDDGWSNKAANVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542  
Qy 1083 FNATSAHFEGEGGPIWDLNCTGESHLCWCPGSRGSGHDCRKHEDAGVICSFTALR 1142  
Db 761 FNATSAHFEGEGGPIWDLNCTGESHLCWCPGSRGSGHDCRKHEDAGVICSFTALR 1142  
Qy 1143 LYSETETSCAGRLVYFNGTWSGVRNITTAIAGVCRQLCGGNGVVSAPLSKTS 1202  
Db 821 LYSETETSCAGRLVYFNGTWSGVRNITTAIAGVCRQLCGGNGVVSAPLSKTS 1202  
Qy 1203 GFVWDDIOCPKTHSIWOCLSAPHERISSPAEETWITCEDRIVRGDCGTCGRVEIW 1262  
Db 881 GFVWDDIOCPKTHSIWOCLSAPHERISSPAEETWITCEDRIVRGDCGTCGRVEIW 1262  
Qy 1263 HAGSWGTVCDDSDWDLAEAEVVCQQLCGSALALDASFGQGTGTIWLDDMRCKNESPL 1322  
Db 941 HAGSWGTVCDDSDWDLAEAEVVCQQLCGSALALDASFGQGTGTIWLDDMRCKNESPL 1322  
Qy 1323 WDCHAKPWGQSCGKHEDAGVCRGOSLK-----SLNASSGHALILSIFGLLLVLF 1376  
Db 1001 WDCHAKPWGQSCGKHEDAGVCRGOSLK-----SLNASSGHALILSIFGLLLVLF 1376  
Qy 1377 I--LFTWCRVQOKHPLRVSTPRRGSLEENLFHEMETCLKREDPHGTSDTTPNHC 1434  
Db 1061 I--LFTWCRVQOKHPLRVSTPRRGSLEENLFHEMETCLKREDPHGTSDTTPNHC 1434  
Qy 1435 EDASDTSLLGV---LPASEATK 1453  
Db 1116 EDASDTSLLGV---LPASEATK 1453

RESULT 2

I38005

M130 antigen precursor, splice form 4 - human

N;Alternate names: M130 antigen cytosolic variant 2

C:Species: Homo sapiens (man)  
C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 17-Mar-2000  
C:Accession: I38005; S36079  
R:Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D.  
Eur. J. Immunol. 23, 2320-2325, 1993  
A:Title: A new macrophage differentiation antigen which is a member of the scavenger  
A:Reference number: I38003; MUID:93380506; PMID:8370408  
A:Accession: I38005  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1156 <LAW>  
A:Cross-references: EMBL:Z22970; NID:g312145; PIDN:CAA80543.1; PID:g312146  
C:Comment: For other splice forms, see PIR:I34004 and PIR:I34006.  
C:Genetics:  
A:Gene: GDB:M130; GDB:MM130  
A:Cross-references: GDB:119369; OMIM:133710; GDB:995898  
A:Map position: 10pter-10qter  
A:Note: the locus has not been clearly established and there is some confusion in the  
C:Superfamily: scavenger receptor cysteine-rich domain homology  
C:Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protei  
F:1-40/Domain: signal sequence #status predicted <SIG>  
F:41-1156/Product: M130 antigen, splice form 4 #status predicted <MAT>  
F:41-1045/Domain: extracellular #status predicted <EXT>  
F:43-147/Domain: scavenger receptor cysteine-rich domain homology <SRC1>  
F:151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC2>  
F:258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>  
F:365-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>  
F:470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC5>  
F:575-678/Domain: scavenger receptor cysteine-rich domain homology <SRC6>  
F:711-814/Domain: scavenger receptor cysteine-rich domain homology <SRC7>  
F:816-920/Domain: scavenger receptor cysteine-rich domain homology <SRC8>  
F:921-1024/Domain: scavenger receptor cysteine-rich domain homology <SRC9>  
F:1046-1066/Domain: transmembrane #status predicted <TRN>  
F:1067-1156/Domain: intracellular #status predicted <INT>  
F:100,118,226,315,440,680,689,762,996/Binding site: carbohydate (Asn) (covalent) #st

Query Match 42.1% Score 3429; DB 2; Length 1156;

Best Local Similarity 56.2% Pred. No. 1,3e-214;

Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

Qy 364 GADLELRADGNSNCGRVEYRIHEQWTTICDQNKWQALVQCKQLGCPVSVFSGRRAK 423

Db 41 GTDRLRLVDGNCNCGRVEYKQVEWGTVCNNGWMEAVSVCNQLGCPVTAIRKAPGAN 100

Qy 424 PSNEARDTWINISICTGNSALWDCTYDGAKRT--CFRRSDAGVICSADKDLRLVGAH 482

Db 101 SSAGSRIWDMHVSRCNSESALWDCKHDGKHSNCTHQDAGVTCDSGNLEMLRTGG 160

Qy 483 SPYGRLEVKYQGEWGTVDHWRWTRNAAVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542

Db 161 NMCGRLEIKFGQGWGTVDHWRWTRNAAVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 220

Qy 543 VSCIGNESINWDCSHSGKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602

Db 221 LTCNGESALNCKHOGKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 280

Qy 603 RWTGTCDDGWSNKAANVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542

Db 281 EWGTCDDGWSNKAANVVCVKGCGKPMHVFQMTYFEKASGPFIWDD 542

Qy 663 RNSGWNDCSHSGKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602

Db 341 KHEWGHKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602

Qy 723 TAEEVCRQLECGSAIRVSRPHTERTLHILMSNGSGTGGASLWDCIRWENKQTAACHLN 782

Db 401 EADVVCRLQCGSALTKSYQVYSKIQTANTWFLSSCNGNETSLWDCIRWENKQTAACHLN 782

Qy 783 MEASLICSARHROPRLVGADMPGSRGVKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602

Db 461 EAKITCSARHROPRLVGADMPGSRGVKHCNCHVREDVITVCSGDATWGLRLVGGNRCGRLEVFYQ 602

Qy 843 SLVSGDHFCKGNGLTWAEKFCQEGSETHALCPVIOHPEDTCIHSREVGVCSTYDRL 902

Db 521 SILGAHFGEGNGOIAEEFQCEGHSHSLCPVAPREGTCSHRDVGVCYSRYEIRL 580  
Qy 903 VNGKSCDGOVEINVLGHWSGLCDTHNDPEDARVLCRLCGTALSTGGYIGERSVRV 962  
Db 581 VNGTPEGRVELTGLGAWGLSHNDIEDAHVLCQKCGVALSTPGGARFGKNGOI 640  
Qy 963 WGRHFHCLGNESLDNCQMTVLGAPPCIHGNTSVICTGSLTQPLFCLANVSDPYLSAV 1022  
Db 641 WRHMFHCTGTGTHGMDCPVLTALGASLCPSEQVAVICSGNSQTLSSCNSSSLGPTPTI 700  
Qy 1023 PEGSALICLEDKRLRLVDGSRACGRVEIYHDGFWGTICDDGWDLSDAHVVVCQKLGCGVA 1082  
Db 701 PEESAVACIESGQLRLVNGGRCACGRVEIYHEGSGWTICDDSDWLSDAHVVVCRLGCGEA 760  
Qy 1083 FNATVSAHFGSGPIWLDLNCCTGTHSLWQCPSSRGNGOHDCHKEKEDAGVICSEFTALR 1142  
Db 761 INATVSAHFGSGTPIWLDLNCCTGTHSLWQCPSSRGNGOHDCHKEKEDAGVICSEFTALR 820  
Qy 1143 LYSETETESACGRLEVFYNGTWGSGVRNITTAIGIVCRLGCGENGVS LAPLSKTGS 1202  
Db 821 LTSEASREACAGRLEVFYNGTWGSGVRNITTAIGIVCRLGCGENGVS LAPLSKTGS 880  
Qy 1203 GFMMVDDIQCPKTHISIQWCLSPAPERRISSPAETWITCEDIRVRGGDTCESGRVEIW 1262  
Db 881 IPMMVNDVQCPKPTLWQCPSPSPWEXRLASPEETWITCEDIRVRGGDTCESGRVEIW 940  
Qy 1263 HGSWGTVCDDSDWLAEEVVCQKLGCGSALALRDASFGOGTGTIWLDDMRCKGNESFL 1322  
Db 941 HGSWGTVCDDSDWLDAAVVCQKLGCGPALKAFKAFEGOGTGTIWLDDMRCKGNESFL 1000  
Qy 1323 WDCHAKPWGSDCGHKEKEDAGVRCGQSLK-----SLNASSGHLALILSSIFGLLLVLF 1376  
Db 1001 WDCPARRGHSECGHKEKEDAAVNCVTDISVQKTPQKATGRSSRQSSFTAVGLVGVLLAIF 1060  
Qy 1377 I--LFLWCRVQKQKHLPLRVSTRRRSLEENLFHEMETCLKRED 1419  
Db 1061 VALFELTKRQROR--LAVSRGENLVHQIYREMNNSCLNADD 1102

## RESULT 3

I38006  
M130 antigen precursor, splice form 1 - human  
N:Alternate names: M130 antigen, extracellular variant  
N:Contains: M130 antigen, splice form 2  
C:Species: Homo sapiens (man)  
C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 05-May-2000  
C:Accession: I38006; I38003; S35768; S36077  
R:Law, S.K.; Micklem, K.J.; Shaw, J.M.; Zhang, X.P.; Dong, Y.; Willis, A.C.; Mason, D.Y.  
Eur. J. Immunol. 23, 2320-2325, 1993  
A:Title: A new macrophage differentiation antigen which is a member of the scavenger rec  
A:Reference number: I38003; MUID:93380506; PMID:8370408  
A:Accession: I38006  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1149 <LAW1>  
A:Cross-references: EMBL:222971; NID:g312147; PIDN:CAA80544.1; PID:g312148  
A:Accession: I38003  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-573, 'R', 608-1149 <LAW2>  
A:Cross-references: EMBL:222968; NID:g312141; PIDN:CAA80541.1; PID:g312142  
C:Comment: For other splice forms, see PIR:I34004 and PIR:I34005.  
C:Genetics:  
A:Gene: GDB:M130; GDB:MM130  
A:Cross-references: GDB:I19369; OMIM:133710; GDB:9959898  
A:Map position: 10pter-10qter  
A:Note: the locus has not been clearly established and there is some confusion in the as  
C:Superfamily: scavenger receptor cysteine-rich domain homology  
C:Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protein  
F:1-40/Domain: signal sequence #status predicted <Sig>  
F:41-1149/Product: M130 antigen, splice form 1 #status predicted <MAT1>  
F:41-1078/Domain: extracellular #status predicted <EXT>

F:41-573,'R',608-1149/Product: M130 antigen, splice form 2 #status predicted <MAT2>  
F:43-147/Domain: scavenger receptor cysteine-rich domain homology <SRC1>  
F:151-254/Domain: scavenger receptor cysteine-rich domain homology <SRC1>  
F:258-361/Domain: scavenger receptor cysteine-rich domain homology <SRC3>  
F:385-468/Domain: scavenger receptor cysteine-rich domain homology <SRC4>  
F:470-573/Domain: scavenger receptor cysteine-rich domain homology <SRC3>  
F:608-711/Domain: scavenger receptor cysteine-rich domain homology <SRC6>  
F:744-847/Domain: scavenger receptor cysteine-rich domain homology <SRC7>  
F:849-954/Domain: scavenger receptor cysteine-rich domain homology <SRC8>  
F:955-1057/Domain: scavenger receptor cysteine-rich domain homology <SRC9>  
F:1079-1095/Domain: transmembrane #status predicted <TRN>  
F:1100-1145/Domain: intracellular #status predicted <INT>  
F:100,118,226,315,440,713,722,795,1029/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 41.7%; Score 3396.5; DB 2; Length 1149;  
Best Local Similarity 54.4%; Pred. No. 1.7e-212;  
Matches 597; Conservative 161; Mismatches 295; Indels 45; Gaps 5;  
Qy 364 CADLELRRLADSGNCSGRVEYRIHBMWMTICDQNNKQALVVKOLGCPFSVFSRRAK 423  
Db 41 CTDXELRLVDGENKCSGRVEYKQVEWCTVCNNGHSMSEAVSICNLGCPPTAKPGWAN 100  
Qy 424 FSNPARDIWNISICTGNESALWDCYDQKAKRT--CFRRSDAGVTCSDKADLDLRLVGAH 482  
Db 101 SSAGSGRIWMDHVSCRGNESALWDCXKHDGKHSNCTHQDAGVTCSDGSLNEMRLTRGG 160  
Qy 483 SPYGRLEVKYOGWGTVCYDHRWSTNNAVVCOLGCKCPMHVFGMTYFKEAGSPIWLDD 542  
Db 161 NMSGRIEIKFOGRGTVCDDNFNDHASICRQLECCSAVSFSGSSNFGEGSGPIWDD 220  
Qy 543 VSCIGNESINWDCRSHGKHCNVHREDIVITCSGDATWGLRLVGGSNRCSGRLEVEYFG 602  
Db 221 LICNGESALNCRKHQGWGKNCDAEDAGVTCSDGADLSRLVDPVTECSGRLEVEFG 280  
Qy 603 RWGTVCDDGWSKAAAVVCSQLDCPSSIIIGMGLGNASTGYKGIWLDVSCDGDSDLSMC 662  
Db 281 EWGTICDDGWSYDAAVACKQLGCPPTAVTAIGRVNASKGFGHIWLDVSCQGHFAVWQC 340  
Qy 663 RNSGWNDCSHSEDEVGVCSDASDMELRLVGGSSRCAGKVEVNVQVAGVILCANGWGN 722  
Db 341 KHHEWKHYCHNEDAGVTCSDGSDLELRRLGGSGRCAGTVEVEIQRLLGKVCDRGWLK 400  
Qy 723 IAEVVCRLGCSAIRVSRPHFTERTLHILMSNSGCTGGEASLWDCIRWEVKOTACHLN 782  
Db 401 EADVVCRLGCSALTKSYQVYKIQATNTWLFSSCNGNETSLWCKNWMQWGLTCDHY 460  
Qy 783 MEASLICSARQPLRVGADMPGSGRVEYKHADTWRSVCDSDSFLAANVLCRELNCDAI 842  
Db 461 EAKITCSAHRPRVLVGDIPCSGRVEYKHGDTWGSICDSDSFLAASVLCRELQCVV 520  
Qy 843 SLSVGDHFGKGLTWAKFCQEGSETHALCPVQHPEDTCIHSRREVGVCSR----- 896  
Db 521 SILGAHFGEGNGOIAEEFQCEGHSHSLCPVAPREGTCSHRDVGVCYSRYEIRL 580  
Qy 897 -----YDVLVNLGKSCQCGQVEINVLGHVNSLCDTHW 929  
Db 581 LIGSVTVKGTGLGSHSLFLKPLCLLPVTEIRLVNKGTPCEGRVELTKLGANGSLCNHW 640  
Qy 930 DPEDARVLCRLCGTALSTGGYIGERSVRVWCHRHRCGLGNESLDDNCQMTVLGAPPC 989  
Db 641 DIEDAHVLCQKCGVALSTPGGARFGKNGOIQIWHMFHCTGTEGHMGDCPVTALGASLC 700  
Qy 990 IHNVTSVICTGSLTQPLFPCLANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRCAGRV 1049  
Db 701 PSEQVAVICSGNSQTLSSCNSSSLGPTPTIPESAVACIESQLRLVNGGRCAGRV 760  
Qy 1050 EYHHDGFWGTICDDGWDLSDAHVVVCQKLGCGVAFNATSAHFEGSGPIWLDLNCCTGTE 1109  
Db 761 EYHEGSGWTICDDSDWLSDAHVVVCRLGCGEAFINATSAHFEGTGTPIWLDKMCNCKE 820  
Qy 1110 SHLWQCPSSRGWQHDCHKEKEDAGVTCSEFTALRLYSETETESACGRLEVFYNGTWGSGVR 1169  
Db 821 SRIWQCHSHGNGQNCRHKEKEDAGVTCSEFTALRLYSETETESACGRLEVFYNGTWGSGVR 880

1170	QY	1170	RNITTAIAGIVCRLQCGEENGVS LAPLSKTGSGFMWVDIOCPKTHISTIWQCLSPAWER	1229
			:::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::  :: :: :: ::	
881	Db	881	SSMSFTTVVVCRLQGCADKGIINPALSIDKAMSTPMWVDNVQCPKPGDTLWQCPSSPWEK	940
			:::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::  :: :: :: ::	
1230	QY	1230	RISSPAEETWITCEDRIIRVRGGDTGECGRVEIWHAGSWGTVCDSDWDLAEAEVVCQLGC	1289
			:::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::  :: :: :: ::	
941	Db	941	RLASPSEETWITCDNKIRLQEGPTSCSGRVEIWHGGSWGTVCDSDWDLDAQVVCQLGC	1000
			:::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::  :: :: :: ::	
1290	QY	1290	GSALAALRDSFGOGTGTIMLDDMRKGNESFLWDCHAKPWGSDCCHKEGACVRCGGOS	1349
			:::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::  :: :: :: ::	
1001	Db	1001	GPALKAFNEAFGGQTGPIINLWVCKNGNESSLWDCPAPRGHSECHGKEDAAVNCITDS	1060
			:::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::  :: :: :: ::	
1350	QY	1350	LK-----SLNASSGHALILUSSTFGLLLVLF1--LFTWCRVQKOKHLPKRVSTRRG	1401
			:::  ::  ::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::	
1061	Db	1061	VOKTPOKATTCRRSSQFFIAVGILGVLLAIFVALFLLTKRRQRQ---LAVSSRGEN	1117
			:::  ::  ::  :: :: ::  ::  ::  :: :: ::  :: :: :: ::	
1402	QY	1402	SLEENLFHEMETCLKRED	1419
			:::  ::  ::  ::  ::  ::	
1118	Db	1118	LVHOIOYREMNCLNADD	1135
			:::  ::  ::  ::  ::  ::	

## RESULT 4

antigen Wc1.1 precursor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 18-Jun-1993 #sequence revision: 19-May-1994 #text\_change 05-Nov-1999  
C:Accession: A46496; I45834; S19133  
R:Wijngaard, P.L.; Metzelaar, M.J.; Machugh, N.D.; Morrison, W.I.; Clevers, H.C.  
J. Immunol. 149, 3273-3277, 1992  
A:Title: Molecular characterization of the Wc1 antigen expressed specifically on bovine  
A:Reference number: A46496; MUID:93056489; PMID:1431105  
A:Accession: A46496  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1436 <W1>  
A:Cross-references: EMBL:X63723; NID:g13; PIDN:CAA45255.1; PID:g14  
A:Experimental source: CDA-C08- gamma delta T lymphocytes  
A:Note: sequence extracted from NCBI backbone (NCBIP:117475)  
R:Wijngaard, P.L.; Machugh, N.D.; Metzelaar, M.J.; Romberg, S.; Bensaid, A.; Pepin, L.;  
J. Immunol. 152, 3476-3482, 1994  
A:Title: Members of the novel Wc1 gene family are differentially expressed on subsets of  
A:Reference number: I45834; MUID:94194107; PMID:7511649  
A:Accession: I45834  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-1436 <W2>  
A:Cross-references: EMBL:X63723; NID:g13; PIDN:CAA45255.1; PID:g14  
C:Genetics: C1  
A:Gene: Wc1.1  
C:Superfamily: scavenger receptor cysteine-rich domain homology  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-1436/Product: Wc1 antigen #status predicted <MAT>  
F:25-131/Domain: scavenger receptor cysteine-rich domain homology <SR01>  
F:132-234/Domain: scavenger receptor cysteine-rich domain homology <SR02>  
F:236-340/Domain: scavenger receptor cysteine-rich domain homology <SR03>  
F:373-476/Domain: scavenger receptor cysteine-rich domain homology <SR04>  
F:478-581/Domain: scavenger receptor cysteine-rich domain homology <SR05>  
F:583-686/Domain: scavenger receptor cysteine-rich domain homology <SR06>  
F:687-789/Domain: scavenger receptor cysteine-rich domain homology <SR07>  
F:791-895/Domain: scavenger receptor cysteine-rich domain homology <SR08>  
F:928-1031/Domain: scavenger receptor cysteine-rich domain homology <SR09>  
F:1033-1136/Domain: scavenger receptor cysteine-rich domain homology <SR10>  
F:1152-1255/Domain: scavenger receptor cysteine-rich domain homology <SR11>

88	QY	QIQCFFSFAM---PRFQAVTRHOKIWIJDDVSCYGNESALWECQH---REWGSNCYHGE	141
89	QY		142
90	QY		143
68	Db	QIQCAGAAIFGPGAYFGPL---GPIWLLXTCEGTSTVSDCEHSNINIKRNDGYNHGR	124
91	QY		125
142	QY	DVGVNCYGEANLGLRLVDGNNSSGRVEVAFQBERWGTICDDGNLNTAAVVCROLGCPSS	201
92	QY		202
125	Db	DAGVVCSG---FVRLAGDGPCSGRVEVHISGEAWIPVSGNFTLATAIICIAELCGGKA	180
93	QY		181
202	QY	FISGVWNSPAVLRIWLLDILCOGNELALWNCRHRRGWGNDCHSHNEDVTLTYCSDSLE	261
94	Db		262
181	Db	VSVLGHSELFRESSAQVAAEFCEGEPPELVCPVPCPGTCHHSGSAQVVCSEVR	240
262	QY	LRLVGGRVNRGRVELKTQGRWGTVCCHHKWNNAAADVVCQLQGGTALHPAGLPHLOSQS	321
95	QY		322
241	Db	L-MTNGSSQCEGOVEEMNITSGWRALCASHANLANVICRLQCGVAISTPGGPHLEVEE	299
96	QY		300
322	QY	DVWLDGVSCSNESFELWDCRHSGTVNFDCILHONDVSVICSDGADLELRLADGNNCSGR	381
97	QY		382
300	Db	DOILTRFHCGSAESFLWSCPVYALGPDGSHGNTASVICSGN-----	342
382	QY	VEVRIHEQWMTICDQNNKNEQALVVCQLQCGPFSVFGSRRAKFSNEARDIWINISCTGN	441
98	QY		442
343	Db	-QIQVLPO-----CNDSSV-----SQPTGSAAS-----	363
442	QY	ESALMDCITYDGKAKTCTPRRSADGIVICSDKADLDRLVGAHSPCYGRLEKYYOGEWGTVC	501
99	QY		502
364	Db	-----EDSAPYCSDSR--QLRLVDDGGPCAGRVEILDGSGMTIC	401
502	QY	HDRWSTRNAAVVCQLQCGCKPMHVFGMTYFKEAASPTIWLDDVSCIGNESINWDEHSGWG	561
100	QY		562
402	Db	DDGWLDLDAVVVCRLQCGEALNATGSAHFGAGSGPILWDLNLNCTGKESHWKPCSRGMG	461
562	QY	KUNCVHREDVIVTCSGDATWGLRLVGGNRCSGRELEVYFQGRMGTICDDGWNKAAAVVC	621
101	QY		622
462	Db	QHNCRHKQDAGVICS--EFLALRWVSEDOQACAGLWVFYNGTWSVCNPMEDITVSTIC	519
622	QY	SQLDPCSSIIGMGLNAST---GYGKIWLDDVSCDGDESDLNSCRNSGWNNDCHSHSD	677
102	QY		678
520	Db	RQLGCGDS---GTLNSVALRREGFRPQWVDRIQCRKTDTSLWQCPDPWNYNSCSPKEE	575
678	QY	VGVICSDASDMELRLVGGSSRCAGKVEVNVQAGVILCANGWGMNTAEVVCROLECGSAI	737
103	QY		738
576	Db	AVIWCADS--QOIRLVDDGGRCSGREVLTDGSGWGTICDORWDLDDARVVCQLQCGEAL	633
738	QY	RVSREPHFETRLHLTMSNSGCTGGEASLWDCIRWEKQTAACHLNMASALICSAHRQPL	797
104	QY		798
634	Db	DATVSFSGTGGSPILWDEVNCRGEESQWVRPCSMGRHQHNCNHQEDAGVICSGF--VRL	691
798	QY	VGADMPGCSGRVEVKHADTWRSVCDSDFLSHAANYLCRELNGCDAILSVSGDHFGKGNGLT	857
105	QY		858
692	Db	AGDGPSCGRVEVHSGEAWTPVSDGNFTLPYTAQVICAELCGKRAVSVLGHMPFRESQGV	751
858	QY	WAEKFCQCEGSETHALCPIVQHPEDTCHHSREVGVVCSRYTDYRLV--NGKSOCDGOVEIN	916
106	QY		917
752	Db	WAEFRCDGGEPELWSCPRVPCPGGTCLHSGAAQVVCVSYTEQLMKNGTSQCEGOVEMK	811
917	QY	VLGHWSGLDTHWDPEDARVLCROLSCGTALSTTGKRYIGERSVRVWGHFRHCLGNESLL	976
107	QY		977
812	Db	ISGRWALCASHSANLANVVCRLQCGCVALSTPRGPHLVEGGDQISTAQFHCSGAESFL	871
977	QY	DNQMTVLGAPPCIHGNTVSVICTGSLTQPLFPCPLANVSDPYLSAVPEGSALICLEDKRL	1036
108	QY		1037
872	Db	WSCPVTALGGPDCSHGNTASVICSNGTHQVLPQCNDFLSQPAGSAAESESSPYCSDSRQL	931
1037	QY	RLVDDGSRCAGRVEIYHDGFWGTICDDGWLDSDAHVVCQKLGCCVAFNATVSAHFEGSG	1096
109	QY		1097
932	Db	RLVDGGGPGCGREVLTDGSGWGTICDDDWDLDDARVVCRLQCGCEALNATGSAHFAGSG	991
1097	QY	PIWLDDNLNCTGESHLMWCPGSRGWGHDCHKEHDAGVICTSEFTALRLYSETETESAGRL	1156
110	QY		1157
992	Db	PIWLDDNLNCTGKESHWKPCSRGWRGHDCHKEHDAGVICTSEFTALRWVS--EDQCGAWL	1049
1157	QY	EVFYNGTWSGVRNNTTAAIGIVCRLQCGCGENGWVSLAPLSKTSKGSGFMWDDIQCPKTH	1216

```
|||||
Db 1050 EVFYNGTWSYCRSPMEDITVSVICRQLGCGDGLSNTSVGLREGSRPRWVLDIOCRKMD 1109
Qy 1217 ISIWQCLSAWERRISPAETWTCB-----DRIVRGGDTSCSRVE 1260
Db 1110 TSLWQCPGPMKYSSCPKEAYISCERRPKSCPTAAACTDREKRLRGDSECSGRVE 1169
Qy 1261 IWHAGSWGTVCCDSDWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKNGES 1320
Db 1170 VWHNGSWGTVCCDSDWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKNGES 1229
Qy 1321 FLMDCHAKPMGQSDCGHKEKEDAGVRCG-----QSLKSLNASSGHLALIL 1364
Db 1230 SLMDCAEPMGQSDCKHEEDAGVRCGVRVTLPTTGTAGRTTNSLPGITSLPGVLCIL 1289
Qy 1365 SSIFGLLLVILFILFWCVRQKHLPLAVSTRRRSLEENLHEMETCLKREDPHGTR 1424
Db 1290 GSLFLVLVLVLTQLRW-RAER-----RALSSYEDALAEAVYEELDYLLTOKEGLGSP 1342
Qy 1425 TS-DDTPNHGCEADSDTSLGLVLPASEATK 1453
Db 1343 QMDTDVDPENYDAAEVPVPGTSPSQGNE 1372
```

## RESULT 5

A59386

sanko - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-2001 #sequence\_revision 31-Dec-2001 #text\_change 31-Dec-2001

C:Accession: A59386

R:sanko, S.

submitted to the Protein Sequence Database, March 2001

A:Reference number: A59386

A:Accession: A59386

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-2403 &lt;SAN&gt;

```
Query Match          36.7%; Score 2990; DB 2: Length 2403;
Best Local Similarity 40.5%; Pred. No. 9.2e-186;
Matches 619; Conservative 177; Mismatches 504; Indels 230; Gaps 23;
```

```
Qy 44 TDLELRVNGDGGCGTVVEYKFGQWGTVCDDGNWTTASTVVKOLGCPFSFAM---FRF 100
Db 230 SSLALRVNGDGRGRVEVLYRGSGTVCDDYWDNDANVVCRLGCGWAMSAPGNAQF 289
Qy 101 GOAVTRGKIWLDDVSCYGNESALWECQHRGWSHNCYHGEDVGVNCYG----- 149
Db 290 GQG---SGPIVLDVRCSGHESYLWSCPNGWLNTHNCGHSEDAGVCSAPLSRPTSPDT 346
Qy 150 -----EAMGLRLVDGNNSCSGRVEVKFQERWGTICDDGNWLNNTAAVVCRLGCG 198
Db 347 WPTSHASTAGPESLALRLVNGDRCGRVEVLYRGSGTVCDDSDWTDNDANVVCRLGCG 406
Qy 199 PSSIFSGVNSPAVLRIWLDLILCOGNELALWNCRRHNGHDCSHNEDVTITCYDS- 257
Db 407 GWATSAPGNARFGQSGPIVLDVRCSGYESYLWSCPNGWLNTHNCGHSEDAGVCSAPLS 466
Qy 258 -----SDLELRVNGDGGCGTVVEYKFGQWGTVCDDGNWTTASTVVKOLGCPFSFAM 304
Db 467 PTITLPASTVGSSESLALRLVNGDRCGRVEVLYRGSGTVCDDSDWTDNDANVVCRLGCG 526
Qy 305 CGTALHFAGLPHLQSGDVVWLDVCSGNSFSLWDCRHSCTVNFDCLOHNDVSVICS-D 363
Db 527 CGWAMLPAGNARFGQSGPIVLDVRCSGHESYLWSCPNGWLNTHNCGHSEDAGVCSAPLS 586
Qy 364 GADLELRVNGDGGCGTVVEYKFGQWGTVCDDGNWTTASTVVKOLGCPFSFAM---FRF 423
Db 587 ESSLALRVNGDGRGRVEVLYRGSGTVCDDSDWTDNDANVVCRLGCGWAMSAPGNAQF 646
Qy 424 PSNEARDIWNISICTGNESALWECQHRGWSHNCYHGEDVGVNCYG----- 469
Db 647 FGQSGGPIVLDVRCSGHESYLWSCPNGWLNTHNCGHSEDAGVCSAPLSRPTSPDTLS 706
```

```
Qy 470 -----DKADLDRLVCAHSPCYGRLEVKYQEGWGTVCHDRWSTRNAAVVC-KOLGC 519
Db 707 TITLPPSTVGSSESLALRLVNGDRCGRVEVLYRGSGTVCDDSDWTDNDANVVCRLGCG 766
Qy 520 GKPMHVFGMTYFKFASGPIWLDVSCYGNESINWDCSHGKGNCHVREDDIVITCS--- 576
Db 767 GWATSAPGNARFGQSGPIVLDVRCSGHESYLWSCPNGWLNTHNCGHSEDAGVCSAPLS 826
Qy 577 -----GATW-----GLRLVGSNRCSGRLEVYFQGRWGTVCDDGNWTTASTVVKOLG 617
Db 827 SRPTSPSTVGSSESLALRLVNGDRCGRVEVLYRGSGTVCDDSDWTDNDANVVCRLGCG 886
Qy 618 AVCSQDLDCPSIIGMGLGNASTVGKTIWDDVSCYGNESINWDCSHGKGNCHVREDDIVITCS 677
Db 887 NVCCRRLLCGWATAPGNARFGQSGPIVLDVRCSGYESYLWSCPNGWLNTHNCGHSEDAGV 946
Qy 678 VGVTCSDA-----SMELRLVGSNRCSGRLEVYFQGRWGTVCDDGNWTTASTVVKOLG 713
Db 947 AGVCSAAHSMSTPSPDTLPTITLPASTVGSSESLALRLVNGDRCGRVEVLYRGSGTVC 1006
Qy 714 LCANGMGNIAEVVCRLECGSAIRVSRPHFTERTLHILMSNGCTGEASLMDICRWE 773
Db 1007 VCDSDWTDNDANVVCRLGCGWAMSAPGNARFGQSGPIVLDVRCSGHESYLWSCPNGW 1066
Qy 774 WKOTACHLNMEASLITCSAHR-QP-----RLVGADMPGCSGRVE 809
Db 1067 WLSHNCGHSEDAGVCSAPLSRPTSPDTWPTSHASTAGSESLALRLVNGDRCGRVE 1126
Qy 810 VKHADTVRSVCDSDVSLHAANVLCRELNCDAISLVSVDHFGKGNGLFWAKFCQEGSET 869
Db 1127 VLYRGSWGTVCCDSDWTDNDANVVCRLGCGWAMSAPGNARFGQSGPIVLDVRCSGHES 1186
Qy 870 HIALCPVQHPEDTCIHSREVGVCVRYTD-----VRLVNG 905
Db 1187 YLWSCPNGWLNTHNCGHSEDAGVCSAPLSRPTSPDTWPTSHASTAGSESLALRLVNG 1246
Qy 906 KSCQDGOVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGGYIGERSVRYVGH 965
Db 1247 GDRQGRVEVLYRGSGTVCDDYWDNDANVVCRLGCGWATAPGNARFGQSGPIVLD 1306
Qy 966 RFHCLGNESLNDNCOMTVLGAPPICHTGNVTVCITGSLTQPLFPCLANVSDPYLSAVPEG 1025
Db 1307 DVRCSGHESYLWSCPNGWLNTHNCGHSEDAGVCSAPLSRPTSPDTWPTSHASTAGSE 1365
Qy 1026 SALICLEDKRLRLVDGSRCAGRVEIYHDFGWTICDDGNWLDSDAHVVCOKLGCQVAFNA 1085
Db 1366 SSL-----ALRVNGDRCGRVEVLYRGSGTVCDDYWDNDANVVCRLGCGWATSA 1419
Qy 1086 TVSAHFEGEGSGPIVLDLNCCTGTETSHLWQCPSRGMOHDCRHKEDAGVCSAPLSRPTSP 1135
Db 1420 PGSARFGQSGPIALDDVRCSGHESYLWSCPNGWLNTHNCGHSEDAGVCSAPLSRPTSP 1479
Qy 1136 -----SEFT-ALRLYSETETESACGRLEVEYNGTWGSGVRNRTTITAGIVC 1181
Db 1480 PDTWPTSRASTAGSESLALRLVNGDRCGRVEVLYRGSGTVCDDYWDNDANVVCRLGCG 1537
Qy 1182 ROLGCGENGVSAPLSK---TGSCFMVVDIQCPTHTISWOCLSAPW----- 1227
Db 1538 ROLGCGENGVSAPLSK---TGSCFMVVDIQCPTHTISWOCLSAPW----- 1227
Qy 1228 -----ERRISSAPEETWIT-----CEDRIRVR---GDDTECSGRVEIHWAGSGMT 1269
Db 1595 GVICSAAQSGTSPRPTDWTNLPALTVGSSESLALRLVNGDRCGRVEVLYRGSGMT 1653
Qy 1270 VCDSDWDLAEAEVVCQQLGCGSALAALRDASFGQGTGTIWLDDMRCKNGESFLMDCHAKP 1329
Db 1654 VCDSDWTDNDANVVCRLGCGWAMSAPGNARFGQSGPIVLDVRCSGHESYLWSCPNGW 1713
Qy 1330 WQSDCGHKEKEDAGVRCGQSLKSLNASSGH 1359
Db 1714 WLTHNCGHSEDAGVCSAPLSRPTSPDTLS 1743
```

## RESULT 6

Tl4893  
scavenger receptor cysteine-rich protein precursor - sea urchin (strongylocentrotus purpurus)  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000  
C:Accession: Tl4893  
R:Pancer, Z.; Rast, J.P.; Davidson, E.H.  
Immunogenetics 49, 773-786, 1999  
A:Title: Origins of immunity: transcription factors and homologs of effector genes of the sea urchin  
A:Reference number: 218253; MUID:99328904; PMID:10398804  
A:Accession: Tl4893  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2153 <PAN>  
A:Cross-references: EMBL:AF064259; NID:g4165052; PID:g4165053; PIDN:AAD08654.1  
C:Genetics:  
A:Gene: SRCR12

Query Match 30.0%; Score 2439.5; DB 2; Length 2153;  
Best Local Similarity 38.2%; Pred. No. 4.3e-150;  
Matches 527; Conservative 197; Mismatches 560; Indels 97; Gaps 22;  
QY 32 LNSCFLLISFN-----GTDLELRVNGDPCSGTVEYVKFGQGWTCVDDGN 78  
DB 113 LSSCAVSSHNCASHSEDAGVWCAGSDGNIRLVGSHSEGRVLYVNMNGTCVDDLD 172  
QY 79 TTASTVVCCKLGC-PFSAMFREFGOAVTRHGKIWLDDVSCYNESALWECQREWGSHNC 137  
DB 173 LNDQVACQRLGLGPAVSADAFEGE---SGSILLDNVACITETISLLCSHPGIESHNC 229  
QY 138 YHGEDVGNVCYGEANLGLRLVDNNSCGRVEYKFOERWGTTCDDGNLNTAAVVCRLG 197  
DB 230 GHSEDAGVTCSSRFSGDIRVVGSGIPTGREGVEYFVNGAMGTVCDDLDINDASVACRLG 289  
QY 198 CPSSFTSSGVNVPVLRPIWLDDILCOGNELALNCRHNGHNDCHNEVDVLTCTYDS 257  
DB 290 FGRTASGAGSYGGSGSIFLDNLACTGAESNLLSCPHNGVSHNGHGEDAGVLCSSS 349  
QY 258 SLELRVGTGTRNCRBELKIQGRWGTGCHHKWNAADVVCYKQLGCGTALHFAFLPHL 317  
DB 350 LSSDIRVVGGRSPTEGREGVEYFVNRANGTVCDDLDVNDVNVACRLGFGRAISAPGASY 409  
QY 318 QSGSDVWLDDVSCSGNESFLMDCRHSHTVNFDCILHNDVSVICSDGA--DLELRADGNS 376  
DB 410 GQSGSILLNLACTGAESILLSCPHNGVSHNCAHGEDAGVSCAPSSQESRVRVLGGLN 469  
QY 377 NCSGRVEYRIHQWWTICDQNKNEQALVVCYKQLGCPFPFVFGSRRAKP--SNEARDIWN 434  
DB 470 NREGVEIFLNQWGTVCDDMGTPDANVVCYKQLGTPSG--GSARSAVFGSGVPILLD 527  
QY 435 SISCTGNESALDCTYDGRKAKRTCFRRSDAGVIC---SDKADLDRLVGAHSPCYGRLEV 491  
DB 528 NVCCSGNERSLELCSNNGVHNCGHQEDASVVCVTGVTTPMGNIRIVGGSPSEGRVEV 587  
QY 492 KYOGENGTVCHDRWSTRNAVVCYKQLGCKPMHVFQMTYFKASGPILWDDVSCITGNESN 551  
DB 588 LVGHRWGTVCDDLDINDANVVCYRELGYSAATSATSSAFGQSGGQILLDLRCSSTESS 647  
QY 552 INDCBSHGKNCVHREDIVVTCSDATWG-----LRLVGSNRCGRLEVYFQGRWG 605  
DB 648 LITCPHGVGVHNCASEDAGVVCA--SITSGPVGVTYMRLVGGQNSRQGRLEISINNQG 706  
QY 606 TVCDDGNWKAAYVVCQLDPSSTIIGMGLGNASTGYKGIWLDDVSCDGEDSLWNSCRNS 665  
DB 707 TVCDDSWINDATVVCYKQLGFSAPVSAFTSAHFGQSGGTIWLDDVSCAGNENSLMDCGHR 766  
QY 666 GWGNDCSHSEDVGVTCSDASD--MELRLVGGSSRCAGKVEVNVQCAVILCANGHGMNI 723  
DB 767 GLGVHNCACHAEDAGVYCI--ASDGLPIRLLAGRSGMEGRVEISLGGDWGTVCDDSGIED 825  
QY 724 AEVVCYKQLGCGSAIRVSRPHFTERTLHILMNSGCTGCGEASLWDCIRWEKQTACHLNN 783

DB 826 AHVVCYKQLGFGPALSANTAAASFGGSGSILMDNVQSCGDEATIAECSHNGIGIHNCGHOE 885  
QY 784 EASLICS-AHRQPRVLVGDAMPSCSRVEYVKHADTWRSVCDSESLHAANVLVRELNCGDAL 842  
DB 886 DAGVVCSSRATGEVRLVGNRSQNEGRVEILISGRWGTVCDDLDWDLRADVLCQLGYGNAI 945  
QY 843 SLSVGDHFGKNGLTWAEKFOCEGSETHLALCPVQHP---EDTCIHSREYGVVCSRYTD 899  
DB 946 SAPHSSFFGPGRGGOILLDDVSCGSEDDILDC---SHPPIGTNNCHSEDAGVCCD--LN 1000  
QY 900 VRLVNGKSCQDQGVENVLGHWSGLCDTHWDPEDARVLCRLQCLSCGTALTSTGGKYIGERS 959  
DB 1001 VRIVNGSRNENGRVEILLHDSWATVCCDDNDDIDATVVCYKQLGFPFSYANVAHAFVGGGS 1060  
QY 960 VRYMGHREPHCLGNESLLDNCOMTVLGPACPIHNGTVSVICTGSLTQPLPCLANVSDPYL 1019  
DB 1061 GEIVLDDVECTGDEVSLTECQHAGLTNNCGHSEDAGVICS-----VNV----- 1104  
QY 1020 SAVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFWGTICDDGWDLSDAHVVCKLGC 1079  
DB 1105 -----RLADGNSPAEGRVEYFDQGWGTVCDDNMGWQNGHVICRAVG 1147  
QY 1080 GVAFNATYSAHFEGSGPIWLDDLNCTGTESHLWOCPSRGWQHDCRHKEDAGVICSEPT 1139  
DB 1148 GKAIATVTRAPRGQGPYILDNINCTGSETSLNCSHSLGTHDCHDSEDAGVICSQPD 1207  
QY 1140 ALRL---YSETETESACRLEVFYNGTWGSGVRRNITTAIAGIVCRLQCGGNGVVS LAP 1196  
DB 1208 TIRLAGRSKYE-----GRVEILQNGAMGTVCDDNWDHDAVAVVRELGYATLTLEATSOA 1262  
QY 1197 LSKTSGFWVDDIOCPKTHISIWOLSLAPWERRISSPAETWITCEDIRVR--GGDTE 1254  
DB 1263 SPQGTGTAILLDDLRCSGREIRLVDPNGHLCQHNCHELDAGVCCQDVIEURLVVGQTP 1322  
QY 1255 CSGRVEIWHAGSWGTVCCDSDWLAEEAVVYCCQLGCGSALAALRDASFQGTGTIWLDDMR 1314  
DB 1323 NAGRVELRYQGGWGTVCDDNWDVDAVAVVRELGPRAIAANNAQAFYGPICGTVLLSDIS 1382  
QY 1315 CKGNESFLWDCHAKPWGSDCKHEDAGVRC-----SQOSLKSLSNASSCHLALISSTF 1368  
DB 1383 CGGQATLFCSPNSGVGLNSCNHGEDAGVYVCMKPVRLSG-----SNNPNEGLEVFMNGLW 1438  
QY 1369 G 1369  
DB 1439 G 1439  
RESULT 7  
Tl8524  
scavenger receptor cysteine-rich protein homolog srcrm2 - Geodia cydonium  
C:Species: Geodia cydonium  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000  
C:Accession: Tl8524  
R:Blumach, B.; Pancer, Z.; Diehl-Seifert, B.; Steffen, R.; Muenkner, J.; Mueller, I.  
J. Cell Sci. 111, 2635-2644, 1998  
A:Title: The putative sponge aggregation receptor: Isolation and characterisation of  
A:Reference number: 218947; MUID:98369060; PMID:9701562  
A:Accession: Tl8524  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2043 <BLU>  
A:Cross-references: EMBL:Y14953; NID:el364818; PID:el364819; PIDN:CAA75175.1  
C:Genetics:  
A:Gene: SRCRM2  
Query Match 27.6%; Score 2246.5; DB 2; Length 2043;  
Best Local Similarity 32.4%; Pred. No. 1.4e-137;  
Matches 514; Conservative 196; Mismatches 588; Indels 287; Gaps 31;  
QY 7 SHWIDFGCCCHONLFSAVVTCILLNSCFLLISFNGTGLELRV-NGDPCSGTVEYKF 65  
DB 114 AWH-SVGSECSHQNNSGVYCT-----DDSAVRLVEGDMNRRVEYVH 156

```
Qy 66 QGOMGTVCDDGNTTASTVVCVKQLGCPFSFAMFRFGQAVTRHGKIMLWDDVSCYNESALW 125
Db 157 ADEGSCICDDYFGFGEAVVICHHELGFSAEDSHGGGTFPOGDGLIWMNDLQCTCHESSIA 216
Qy 126 ECQREHNGSHNYGEGDVGVNCEAN-----LGLRLVDGNNSCS-----GRVEYKQ 173
Db 217 ECTFPFGGEHNCQREDAGVCTPATDHPFYIPIRLNNGTSSNATHGLQYGRVEFIN 276
Qy 174 ERMTICDDGNLNTAAVVCRCPCSPSSISSGVVNSPAVLN-----PIWLDDILCOGNELA 230
Db 277 NTWGTVCDDGNGIEDATACRLQFYGALLASSRSEVAQVAPSPYIWDVVLCHGDEVA 336
Qy 231 LWNCRHGMNHDCHSHNEDVTLTCYDSDLE-----LRLVGGTNRCM-----GRVEL 277
Db 337 LSECHHRTTGVHDCSHVEDAGVFC--TTNDIQAPAKVRLVRNNGNSVDEDCGHVEGRVEV 395
Qy 278 KIQRWGTVCHHKNNNAADVVCVKQLGCGTALHFAGLPHLOGSDVVDVWLVGVSNGNSP 337
Db 396 NFYGVWGTICDDYWSLADGVICKMLGFTYADEVGYSGAFGIGTGPIWMDVDDVQCDGTET 455
Qy 338 LWCDRHSGTVNFDCLHONDVSVICS---DGADLELRADGNNSGRVEVRIHEQWWTIC 394
Db 456 IDHCDFFGNGVHNYCHPEDASVVCYRQPGSD--VRLVGGSSMEGRVEVLNGVWGTVC 513
Qy 395 DONKNNEOALVVCVKQLGCPFSFVSGRRAP-----SNEARDIWNISICTGNESALWDC 448
Db 514 DDSNGAEDAQVICHQLG----FYGTATAVRGTQGFASSSQPIWLDDVGCTGSEMYLSDC 569
Qy 449 TYDGKAKRTCFRRSDAGVIC-----468
Db 570 SSSGFGNNHNCYHEDAGVGVCOGTSAFIPVRLNSNTRPNEGTVEIFYQNRWGTICDIFWTL 629
Qy 469 -----468
Db 630 YDANVVCRLQYDGNAYNATHSSYCGDSLTTTAYTDCFGQESQLANCTGPRYPPIPSW 689
Qy 469 -----SDKADLRLVGAHSPCYGRLEVYQGEWGTVCHDRWSTNAAVVC 514
Db 690 YCGDNTVAGVMCIGTEDGPAPIRLVGSEASEGRVEILVQGWGTICDDIWDQADAEEVC 749
Qy 515 KQLG-CGKPMHVFMTYFKEASGPTLWDDVSCIGNESINWDCSHGKGNKNCVHREDVIV 573
Db 750 RQLGYLSEGATALQFAHFGESGPIVDDVNCGTGLELYTDCPSSGFIYHNCASEDAGV 809
Qy 574 TCSG-----DATWGLRLVGGSS---NRCSGRLEVYFQGRWGTVCDDGWNKAAVVCSSOLD 625
Db 810 RCNARGPQOTYPIRLVRNGTVITQNEGTVIELHNGTWSAVCDDYNGYTEAVVACHML- 868
Qy 626 CPSSIIIGMLGNASTGY-----GKIWLDDVSCDGEDSLWSCRNSGNGNDCSHSE 676
Db 869 -----GFATAVRATRSRPHGAVDGTFLDNVRCACGTEREIDFVHSSYSRNCDRTO 920
Qy 677 DVGVTCSO--ASDMELRLVGGSSRCAGKVEYNVOGAVGILCANGWGMNIAEVVCRLBEG 734
Db 921 QAGVACTNYTASEPIRLVGGSGPHEGRVEIYOGVWGTVCDDGWSQPDADVVCRLQGYA 980
Qy 735 SAIRVSREPHEPTERTLHILMNSGTCGEASLWDCIRNWEKQTACHLNEASLICSASRHQ 794
Db 981 NASRATVRAEFGRTGETIWLNVACTGFENSLDECNSGNGDHNCGHREDAGAVCOGELL 1040
Qy 795 P-RLVGADMPSCGRVEVKHADTRWSVCDSDFSLHAANYLVCRELNCGDAISLVGDHFKG 853
Db 1041 PIRLRDGSNLEGRVEVYFNKWTGTCVDDFDWLDGATVVCRLQGYPEAVRAEPFAFGSG 1100
Qy 854 NGLTWAKEFCQEGSETHALCPVQHEDTCHHSREVGVVC-----SRYTIDVRLVNGKS 907
Db 1101 EGPVWLDDVACVGTETSLFECSSGLGQHNCOHFEDAGVECDNHTTTPSPVTDLRAGGTN 1160
Qy 908 QCDGOVENLWGHGSLCDTWDPEDARVLCRLQSCGTALTSTGKTYI--GERSVRVNGH 965
Db 1161 ATEGRVEIFFNNTWGTICDDSDWIDRAEVVCRYLGFESAIALSNFYGAGDPOPIWLD 1220
Qy 966 RFHCLGNESSLNDNCQMTVLGAPPCIHGNTVSVICTGSLTQPLFPLCLANVSDPYLSAVPEG 1025
```

```
Db 1221 DVDCFGSSTITSCLTSALEHNCARHVEDAGVRC--VISHPL-----1260
Qy 1026 SALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGWDLSDAHVVCOKLGCQVAFNA 1085
Db 1261 -----LRLVNGMTGEGRVEVPHRGSWGTVCDDHSEIDANIVCKELGFARALISA 1310
Qy 1086 TVSAHFEGEGSGPIWLDDLNCVTGTSHELWQCPSRSGQWDCRHKEDAGVICSEF-----TAL 1141
Db 1311 SGFATFEGEGSGTIWLDDVQCDGTEATIFNCFSSGNGNHNCLHSEDAISAICTAIPVNPV 1370
Qy 1142 RLYSETETESACGRLEVYNGTWSGVRNNTTAAGTIVCOLGC-GENGVVSLAPLSKT 1200
Db 1371 QLIDGPSVHE--GRQIYYNNWGTVCDDQWGYDEANYVCKSLGPGADPNLSLLNYP 1428
Qy 1201 GSGFMWDDDIQCPKTHISIMOCLSAPWERRISSPAETWITCEDR---IRVRGGDTECSG 1257
Db 1429 GSGHILWDDVDRCRDEFFIQDNCNADIGENNCGHYEDVGLRCLPNTLVDRLMNGNTSAG 1488
Qy 1258 RVEIWHAG-SWGTVCDDSDWDLAAEAEVVCQQLGCGSALAALRADSAFGQGTGTIWLDDMRCK 1316
Db 1489 RVEVYNGEENGWGTVCDDNDWDIKDADVVCRLMDFKSAQAPREAFEGDGLGVWEDDFLCS 1548
Qy 1317 GNEFLWDCCHAKPGQSDCGHKEKEDAGVRCQSLKSLNASSGHLALILSSIFGLLLVLF 1376
Db 1549 GYEDSLLECSHAGVKVHNCRHSEDAVSVC-----1578
Qy 1377 ILFTWCRVQKQKHLPLRVSTRRRG 1401
Db 1579 --MLTVCRDFSQEHGTFOISNRQFG 1601

RESULT 8
T30549
hensin - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30549
R;Takito, J.; Hikita, C.; Al-Awqati, Q.
J. Clin. Invest. 98, 2324-2331, 1996
A;Title: Hensin, a new collecting duct protein involved in the in vitro plasticity of
A;Reference number: 220851; MUID: 97096804; PMID: 8941650
A;Accession: T30549
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1594 <TAK>
A;Cross-references: EMBL:AF043112; NID: g4105083; PID: g4105084; PIDN: AAD02242.1

Query Match 26.5%; Score 2156.5; DB 2; Length 1594;
Best Local Similarity 36.6%; Pred. No. 7.1e-132;
Matches 475; Conservative 132; Mismatches 387; Indels 303; Gaps 29;

Qy 46 LELRLVNGDGPCSGTVEVKFQOGNGTVCDDGNTTASTVVCVKQLGCPFSFA---NFRFGQ 102
Db 51 LELRLANGDRCQGRVEVLYQGSWGTVCDDGWDINDAQVVCRLQCGMAVSAAPSAREFG 110
Qy 103 AVTRHGKTLWDDVSCYGNESALWECQREHNGSHNYGEGDVGVNCEY-----149
Db 111 G---PGQIVLDDVSCSGQEPYLMWSCHHRGWSLHNGHCHQEDAGVICSDAMAMTSPPTDTP 167
Qy 150 -----EANLGLRLVDGNNSGRVEVYFQERWGTICDDGWNLTAA 190
Db 168 TTVIYESTPHFPFSGLELVFPSCGLELRLANGSDRCGRVEVLYQGSWGTVCDDGWDINDAQ 227
Qy 191 VVCRLQCGPSSFISSGVVNSPAVLN-----PIWLDDILCOGNELANCRHGRGNHDC 244
Db 228 VVCRLQCGMA-----VSAPGARFGQPGQIWLDDVSCSQEPYLMWSCHHRGWSLHNC 281
Qy 245 SHNEDVTLCYDSD-----SDLELRVLGGTNRCMGRVVELKIQ 280
Db 282 QHEDAGVICSDAVPTTTPPDTPVTIYESSVPFSGLELRLANGSDRCGRVEVLYQ 341
Qy 281 GRWGTVCHHKNNNAADVVCVKQLGCGTALHFAGLPHLOGSDVVDVWLVGVSNGNSFLWD 340
```



Db 796 GQGTPIVMDVACGGYEDYLWRCSHRWLSHNCGHEDAGVICSASQSSPTDPVFYPT 855  
QY 686 -----SDMELRLVGGSSRCAGKVEVNVQAVGILCANGMGMTAEV 726  
Db 856 DQTVETQVTPYPTGTENSLAVRLNGDRCQGRVEILYQSGWTVCDSDWTDKDV 915  
QY 727 VCRQLCGSAIRVSRPHEPFTLHLMSNGSGTGGEASLWDCIRWEWKOTACHLNMEAS 786  
Db 916 VCRQLCGWAVSAPGNAYFGPGSGVILDDVACTGHEDYLWCSHRWLSHNCGHEDAG 975  
QY 787 LIC-----AHROPRL-----VGADMPGSG----- 806  
Db 976 VICSDAQIOSTTRPDLWPTTTTETTELTTTPYDFDWMWTTSDYSCGGLLTOPSGQFSS 1035  
QY 807 -----RVEYKHADTW-----RSVC 820  
Db 1036 PYPSYNNARCWSKIVLPNMNRVTVFTDVLQEGGCNDYILVYDGPENSSLIARVC 1095  
QY 821 DSD-----FSLHAANVLCRELNGDGLSLVSGDHFGKNG 855  
Db 1096 DGSNGSFTSTGNFMSVVFTDGSVTRRGFOAHYYSTVSTNYSCGGLLTOPSGQFSSPYYP 1155  
QY 856 LWAERFQCGSETHLALCPVHPEDTCHSREVGVCVSRVTDVRLVNCCKSQ----- 908  
Db 1156 SNYPNNARCWE-----ILVP-----NMNRVTV-----FTDVLQEGGCNDYILVYD 1198  
QY 909 -----CDGQVEINVLHGWSLCTDHPEDARVLCRLQSCCTALSTTGKRYIG 956  
Db 1199 GPQYNSSLIARVCDGSN-----GSFTST-----GNFMSVV-----FIT 1231  
QY 957 ERSVRVWGHFRHCLGNESLNDNCOMTVLGAPPCHGNMTVSVICTGSLTOP----- 1006  
Db 1232 DGSVTRRGFOAHYYS-----TVSTTPVPVPTDSDYSCGGLLTPSGQFSSPHY 1281  
QY 1007 -----LFCPLANVSDPYLSAVDEGSALICLEDKRLVLDGDSRCAGRVEIYH 1053  
Db 1282 SNYPNNARCWEILVPMNRVTVFTDVLQEGG-----CNYDYIL-VYDGE-----YN 1329  
QY 1054 DGFWTGICDDGDLSDAHVVCQKLGCVAFNATVSAHFGEVGPVWLDLNLCT--GPESH 1111  
Db 1330 SLLIARVCD-----GNSGFTST-----GNFMSVVFTDGSVTRRGFOAH 1369  
QY 1112 LMQCPSRGW-----QOHDCH-----KEDAGVICS----- 1136  
Db 1370 YYSTVSTNYSCGGLLTOPSGQFSSPHYSPYNNPNNVRCSEILVPSMNRVTVFTDVLQEG 1429  
QY 1137 --EFTALRLYSEFETES-----CAGLEVFYNGTWGCVGRNITTAIAGIVCRQLCCGE 1188  
Db 1430 GCSFDYILVYDGPENSSLIARVCDG-----FNGSFTSTG--NFMS----- 1468  
QY 1189 NGVYSLAPLSKTCGFMWVDIQCPRKTHSIWQCLSAPEWERRISSPAETWITCED--R 1245  
Db 1469 --VFTDGSVTRGF-----QAYYSTVSTSLHHNITGNDSLSA 1507  
QY 1246 IRRVGGDTECSGRVEIWHAGSWGTVCDSDWDLAEAYVCOQLCGGSALALRDSAFQGT 1305  
Db 1508 LRLVNGSNRCRGRVEILYRSGWTVCDSDWGISDANVVCRLCGGSALSAPGNWFGQGS 1567  
QY 1306 GTTWLDMCKGNESFLWDCNAPWGQSDCGHEDAGVICS 1346  
Db 1568 GLVILDDVSCSGYESHLWNCHHPGLVHNCRHSEDAGVICA 1608

## RESULT 10

S56744

mucin (clone pgM7-1) - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 27-Oct-1995 #sequence\_revision 30-Jan-1998 #text\_change 20-Apr-2000

C:Accession: S56744

R:Nunes, D.P.; Keates, A.C.; Afdhal, N.H.; Offner, G.D.

Biochem. J. 310, 41-48, 1995

A:Title: Bovine gall-bladder mucin contains two distinct tandem repeating sequences: ev

A:Reference number: S56744; MUID:95374471; PMID:7646470

A:Accession: S56744

A:Molecule type: mRNA

A:Residues: 1-600 &lt;NUN&gt;

A:Cross-references: GB:S78981; NID:g1042036; PIDN:AAB35069.1; PID:g1042037

C:Superfamily: scavenger receptor cysteine-rich domain homology

F;1-102/Domain: scavenger receptor cysteine-rich domain homology (fragment) &lt;SRC1&gt;

F;126-229/Domain: scavenger receptor cysteine-rich domain homology &lt;SRC2&gt;

F;253-356/Domain: scavenger receptor cysteine-rich domain homology &lt;SRC3&gt;

F;380-483/Domain: scavenger receptor cysteine-rich domain homology &lt;SRC4&gt;

Query Match 13.2%; Score 1076.5; DB 2; Length 600;

Best Local Similarity 38.5%; Pred. No. 2.9e-62;

Matches 223; Conservative 52; Mismatches 157; Indels 147; Gaps 13;

QY 48 LRLVNGDPCSGTVEKFGQWGTVCDDGNWNTTASTVVKQKCPFPFAM---FRFGQAV 104

Db 2 LRLVNGSDRCQGRVEVLYGGWGTVCDDSDWTDNANVVCRLGCGWALSAPGDARFGQ- 60

QY 105 TRHGKIWLDDVSCYGNESALWECQHRWGSNCHYGHEDGVNVC----- 147

Db 61 --SGPIVLDVCGSGYETYLWCSHSPWNTHNCHGHSEDAVICSASQTQSTVVPDWLYPT 118

QY 148 --YG-EANLGLRLVDGNNSCSGRVEVKEFQERWGTICDDGNLNTAAVVCRLGCPSPFIS 204

Db 119 TDYGTESGLALRLVNGDRCQGRVEVLYRGSWGTVCDDSDWTDNANVVCRLGCG- 172

QY 205 SGVYNS-PAVLR-----PIWLDDILCOGNELALWNCRHRCGNHDCSHNEDVTLTC--- 254

Db 173 -GWANSAPGNARFGQSGPIVLDVCGSGYETYLWCSHSPWNTHNCHGHSEDAVICSAS 231

QY 255 -----YDSSD-----LELRLVGGTNRCMGRVELKIQGRWGTVCVHHKWNNAAD 297

Db 232 QTQSTVVPDLWYPTDYGTESGLALRLVNGSDRCQGRVEVLYRGSWGTVCDDSDWTDNAN 291

QY 298 VVKQLCCGTALHPAGLPHLQSGSDVVLVGVSCSGNESFLWDCRHSCTVNFDCQLHQNDV 357

Db 292 VVCRQLCGWGISAPGDARFGQSGPIVLDVCGSGYETYLWCSHSPWNTHNCHGHSEDA 351

QY 358 SVICSDGADLELRLADGNSNCSGRVEVRIHEQWNTICDQNNKNEQALVVCQKCPESVF 417

Db 352 SVICSA-----SQTQSTV----- 365

QY 418 GSRAKPSNEARDIWIINSICTGNESALWDCITYDGAKARTCFRRSDAGVICSADKADLDR 477

Db 366 -----PDLW-----YPTDYG-----ESGLALR 384

QY 478 LVGAHSPCYGRLEVKYQGEWGTVCCHDRWSTRNAAVVCQKCGCPMHVFGTYKEASGP 537

Db 385 LVNGSDRCQGRVEVLYGGWGTVCDDSDWTDNANVVCRLGCGSGISAPGDARFGQSGP 444

QY 538 IWLDDVSCIGNESNIWDCSHGKGNCHVHREDVIVTCS 576

Db 445 ILLDDVCGSGYETYLWCSHSPWNTHNCHGHSEDAVICS 483

## RESULT 11

A57190

ebnerin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 05-Nov-1999

C:Accession: A57190

R:Li, X.J.; Snyder, S.H.

J. Biol. Chem. 270, 17674-17679, 1995

A:Title: Molecular cloning of Ebnerin, a von Ebner's gland protein associated with ta

A:Reference number: A57190; MUID:95353532; PMID:7629065

A:Accession: A57190

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-1290 &lt;LIA&gt;

C:Cross-references: GB:032681; NID:g975346; PIDN:AAC52248.1; PID:g975347

C:Superfamily: scavenger receptor cysteine-rich domain homology; C1r/C1s repeat homol

C:Keywords: extracellular protein; transmembrane protein

F;1-77/Domain: scavenger receptor cysteine-rich domain homology #status atypical &lt;SRC

F:142-245/Domain: scavenger receptor cysteine-rich domain homology <SRC1>  
F:289-392/Domain: scavenger receptor cysteine-rich domain homology <SRC2>  
F:583-689/Domain: C1r/C1s repeat homology <C1R1>  
F:713-816/Domain: scavenger receptor cysteine-rich domain homology <SRC3>  
F:838-944/Domain: C1r/C1s repeat homology <C1R2>

Query Match 12.98; Score 1048.5; DB 2; Length 1290;  
Best Local Similarity 29.78; Pred. No. 4.5e-60;  
Matches 289; Conservative 114; Mismatches 290; Indels 281; Gaps 34;

QY 286 VCHKWNAAADVCKQLGCGCTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWDCRHS 345  
DB 1 MCDSDWINDANVCRQLGCGWALSAPGSAQFGGLGPIVLDVACRGHEAYLWSCSRG 60  
QY 346 TVNFCDLQNDVSVICSDGADLELRLADGNNCSGRVEVRHEQWTTICDQWNEQALV 405  
DB 61 WLSHNCGHQEDAGVICS-----SQTSSPTPG-----WW-----NPGGTNDVIY 100  
QY 406 VKQLGCGCFVSRRAPKPSNEARDIWNISCTGNESALWDCYTDGKAKTCFRSDAG 465  
DB 101 DTQET-----TETSQTSSTPDW-----NHGTTINDVIYD---TQETTEGTDG 142  
QY 466 VICSDKADLLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRAAVVCKQLGCGKPMHV 525  
DB 143 -----LAVRLVNGDRGRVEILYQSGVIVCDSDWINDANVCRQLGCGWALS 194  
QY 526 FGMYFKASGPIWLDDVSCIGNESNIWDCSHGSKHNCVHREDVIYVCS----- 576  
DB 195 PGSAGFGGSGSIVLDDVACRGHEAYLWSCSRHGLSHNCGHQEDAGVICSQTSPTP 254  
QY 577 -----GDATWG-----LRLVGGNRCGRLEVFQGRWG 605  
DB 255 DSQTSPTPGWNNPGGTNDVSYGPEQTTDATDGLAVRLVNGDRGRVEILYQSGW 314  
QY 606 TVCDDGNSKAAAVVCSLDPCSSIIIGLGNASTGYKIKLDDVSCDGEDSDLSWCSRNS 665  
DB 315 TVCDDSDTKDANVCRQLVCGWALSAPGSAHFGGSGSIVLDDVACRGHEAYLWSCSR 374  
QY 666 GWNNDCHSDEVDGVICSAS-----DM-----ELRLVGGSRCAKGVENVQGVAG 712  
DB 375 GWLSHNCGHEDAGVICSDAQTSTWPDWMTPTTPTTDDWTTTSSSYPTTQFPPI- 433  
QY 713 ILCANGWGNIAEVVCRQLGCGSALRVS-----REPHEFERTLHILMSNGCTGGEASLWD 768  
DB 434 ---ADWTTTPPEY-----CGLLTLPLYGQFSSYPGCS-----YPNARC-----LW- 474  
QY 769 CIRWEKQTAACHLNEASLICSASRQRLVADMPGCSGRVEYKADTWRSVCDSDFL-- 826  
DB 475 -----KIFYSSMNRVTVVDVQLEGG-----CNYDIILVF 505  
QY 827 ---HAANVLCRELNGDAISLVGDHFGK-----NGLTWAEKFOCE--- 865  
DB 506 DGPENSSLIAR-----VCDGFGSFTSTQNFMSVFIIDGVSVTRRGFOADYYS 554  
QY 866 ---GSETHLALCPV-----QHPEDTC--IHSREVGVSRYTDDVRLVNSKSCDQGV 913  
DB 555 TPISTSTSTPTTFFIVTDWTTTPPEYTCGGLLTLPLYGQFSSPY----- 598  
QY 914 EINVLHGWSLCTDHWDPEDARVLCR-----QLSCGT-----ALSTTGK 953  
DB 599 -----YFGSY-----PNNARLWKIFVPSMNRVTVVDVQLEGGCNYDIILGFDGPE 646  
QY 954 YIGERSVRVWGHFRHFCNLGNSLLDNQMTVLGAPPCIHGNVTSV--ICTGSLTQPLF--- 1008  
DB 647 YNSSLIARV-----CDGNSGFTSTQ-----NFMSSVFIIDGVSVTRRGFOAD 688  
QY 1009 ---PCLANSDPYLSAVPEG--SALICLEDKRLRLVDGSRCAKGVREIYHDFGWTICD 1062  
DB 689 YVSTPIRSTTPTTPTTFFITGNDSSLV-----LRLVNGTNRCEGRVEILYRGSWVPCAD 742  
QY 1063 DGWDLSDAHVVCQKLGCGVAFNATYSAHFGESGPIWLDDNLCTGTESHLWQCPGRWGQ 1122  
DB 743 DSWDINDANVCRQLGCGSALSAPGNWFGGSGGLIVLDDVSCSGYESHLLNCRHPGLV 802

QY 1123 HDCRHKEDAGVICS 1136  
DB 803 HNCRRVEDAGVICS 816

## RESULT 12

TI7405  
C:Species: Strongylocentrotus purpuratus (purple urchin)  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000  
C:Accession: T17405  
R:Pancer, Z.; Rast, J.P.; Davidson, E.H.  
Immunogenetics 49, 773-786, 1999  
A:Title: Origins of immunity: transcription factors and homologs of effector genes of  
A:Reference number: Z18253; MUID:99328904; PMID:10398804  
A:Accession: T17405  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1036 <PAN>  
A:Cross-references: EMBL:AF076513; NID:g4164530; PID:g4164531; PIDN:AA05493.1

Query Match 10.68; Score 866.5; DB 2; Length 1036;  
Best Local Similarity 27.18; Pred. No. 2.3e-48;  
Matches 229; Conservative 80; Mismatches 227; Indels 309; Gaps 18;

QY 312 AGLPHLQSGSDVWLDGVSCSGNESFLWDCRHSHTVNFDCIHLQNDVSVICSDGADLELRL 371  
DB 478 SGIPAL-----LDNIAADVDSIL-----SVLTDA---QOCSTIVQEGDLALTY 519

QY 372 ADGSNNCSGRVEVRHEQWTTICDQWNEQALV-----CKQLGCFPSVFGSRR 421  
DB 520 IPGLD-----PYALNLQIYQITGMNGTNQLCQVVADAFQPGDHHCPD-GMTFNECGSGC 572

QY 422 AKPSNE---ARDI---WINSISCTGNESALWDCYDGKAKRTCRRSDAGVIC---SDKADL 474  
DB 573 GPGSCDNLVPRDICTPLFCFVCGCFEGLVKD-----QDGDRCIPVDQCG 618

QY 475 DLRLVGAHSPCYGRLEVKYQGEWGTVCHDRWSTRAAVVCKQLG-----CGKPMHVGMTY 530  
DB 619 AIRLVGGNEAGREVEIQYNGVMGTICDSSWGITDASVVCRLMGFGQAGAP---GSAH 674

QY 531 FKEASGPIWLDDVSCIGNESNIWDCSHGSKHNCVHREDVIYTCSDATWGLRLVGGSN 590  
DB 675 FQGGTPIQLDDVCGTGAEQITFDCAHPAFGVHNCACHEDAGVFCI---ASQDVRVGGSN 732

QY 591 RCSGRLEVIYFQGRWGTCDGGMNSKAAAVVCSQLDPCSSIIGMGLGNASTGYGKIWLDDV 650  
DB 733 EAEGRVEIQYNGVMGTICDSSWGITDASVVCRLMGFGQAGAPGSAQFGQGTGLIQLDDV 792

QY 651 SCDDGDESILWSCRNGWNNDCSHEDVGVICSADSMELRLVGGSSRCAGKGVENVQGA 710  
DB 793 GCTGAEQITFDCAHPAFGVHNCACHEDAGVICSQD---VRLVGGSEAGREVEIQYNGV 850

QY 711 VGILCANGWGNIAEVVCRQLGCGSALRVSREPHEFERTLHILMSNGCTGGEASLWDCI 770  
DB 851 WGTICDSSWGITDANVCRQLMGFGQAGAPGSAHFGQGTGPIQLDDVCGTGVEQITFDC- 909

QY 771 RWEWKQTAACHLNEASLICSASRQRLVADMPGCSGRVEYKADTWRSVCDSDFLHAAN 830  
DB 910 -----AH-----PPFGVH--- 917

QY 831 VLCRELNGDAISLSVGDHFGKGNLTWAEKFOCEGSETHLALCPVQHPECTCHSREV 890  
DB 918 -----NCAHYEDA 925

QY 891 GVCSRYTDVRLVNGKSQCDGQVEINVLHGWSLCTDHWDPEDARVLCRQLSCGTALSTT 950  
DB 926 GVVCILVSQDV----- 935

QY 951 GCKYIGERSVRVWGHFRHFCNLGNSLLDNQMTVLGAPPCIHGNVTSVICTGSLTQPLPFC 1010  
DB 936 ----- 935





GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 05:30:12 ; Search time 21 Seconds  
(without alignments)  
2869.767 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDFGRCCHQ.....CEDASDTSLGLVLPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3082	37.9	1436	1	WC11_BOVIN
2	822.5	10.1	875	1	NETR_HUMAN
3	719.5	8.8	761	1	NETR_MOUSE
4	691.5	8.5	347	1	CD5L_HUMAN
5	664.5	8.2	352	1	Q9GK4_MOUSE
6	633.5	7.8	757	1	LOL4_MOUSE
7	629	7.7	754	1	LOL3_MOUSE
8	622.5	7.6	753	1	LOL3_HUMAN
9	618	7.6	756	1	LOL4_HUMAN
10	614.5	7.6	665	1	CD6_MOUSE
11	607.5	7.5	774	1	LOL2_HUMAN
12	592	7.3	668	1	CD6_HUMAN
13	520	6.4	532	1	SPER_STRPU
14	298	3.7	483	1	MRCO_MESAU
15	290.5	3.6	518	1	MTCO_MOUSE
16	290	3.6	458	1	MSRE_MOUSE
17	288.5	3.5	520	1	MRCO_HUMAN
18	288	3.5	451	1	MSRE_HUMAN
19	279.5	3.4	453	1	MSRE_BOVIN
20	279.5	3.4	454	1	MSRE_RABIT
21	264	3.2	494	1	CD5_MOUSE
22	258.5	3.2	2470	1	NTC2_MOUSE
23	258	3.2	5376	1	2AN_MOUSE
24	254.5	3.1	495	1	CD5_HUMAN
25	248.5	3.1	2471	1	NTC2_HUMAN
26	247.5	3.0	2471	1	NTC2_RAT
27	246.5	3.0	2531	1	NTC1_MOUSE
28	244.5	3.0	2871	1	FBN1_MOUSE
29	242.5	3.0	491	1	CD5_RAT
30	240.5	3.0	1238	1	JAG2_HUMAN
31	239.5	2.9	4544	1	LRP1_HUMAN
32	239	2.9	2531	1	NTC1_RAT
33	233	2.9	1700	1	BAR3_CHITE

RESULT 1

ID	WC11_BOVIN	STANDARD;	PRT;	1436 AA.
AC	P30205;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Antigen WC1.1 precursor.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID:9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood;			
RA	MEDLINE=93056489; PubMed=1431105;			
RA	Wijngaard P.D.J.; Metzelaar M.J.; Machugh N.D.; Morrison W.I.;			
RA	Clevers H.C.;			
RT	"Molecular characterization of the WC1 antigen expressed specifically on bovine CD4-CD8-gamma delta T lymphocytes.";			
RL	J. Immunol. 149:3273-3277(1992).			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- SIMILARITY: CONTAINS 11 SRCR DOMAINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL: X63723; CAA45255.1; -			
DR	PIR: S19913; S19913.			
DR	PIR: A46496; A46496.			
DR	InterPro: IPR001190; Srcr_receptor.			
DR	Pfam: PF00530; SRCR; 11.			
DR	PRINTS; PR00258; SPERACTRCPTR.			
DR	SMART: SM00202; SR: 11.			
DR	PROSITE; PS00420; SRCR_1; 4.			
DR	PROSITE; PS0287; SRCR_2; 11.			
KW	Antigen; Repeat; Glycoprotein; Signal.			
FT	SIGNAL 1 25 POTENTIAL.			
FT	CHAIN 26 1436 ANTIGEN WC1.1.			
FT	DOMAIN 28 131 SRCR 1.			
FT	DOMAIN 134 234 SRCR 2.			
FT	DOMAIN 239 340 SRCR 3.			
FT	DOMAIN 376 476 SRCR 4.			
FT	DOMAIN 481 581 SRCR 5.			
FT	DOMAIN 586 686 SRCR 6.			
FT	DOMAIN 689 789 SRCR 7.			
FT	DOMAIN 794 895 SRCR 8.			
FT	DOMAIN 931 1031 SRCR 9.			
FT	DOMAIN 1036 1136 SRCR 10.			
FT	DOMAIN 1155 1255 SRCR 11.			

P35555 homo sapien  
P46531 homo sapien  
P07207 drosophila  
P98133 bos taurus  
P97607 rattus norv  
Q9tv36 sus scrofa  
P61555 mus musculus  
P19238 bos taurus  
P35556 homo sapien  
P98157 gallus gall  
Q61982 mus musculus  
Q9qy65 mus musculus

ALIGNMENTS

FT CARBOHYD 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 244 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 256 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 351 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 424 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 444 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 531 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 717 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 799 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 897 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 979 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 999 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1086 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1173 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1393 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1436 AA; 154196 MW; 6C72E5FDC6BA088C CRC64;  
 Query Match 37.9%; Score 3082; DB 1; Length 1436;  
 Best Local Similarity 41.7%; Pred. No. 1.2e-208;  
 Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;  
 QY 28 CILLNSCEPLISFNGTDLRLVAGDPCSGTVEYKQFGQWGTVCDDGWNNTASTVYVK 87  
 Db 13 CVLLG-----TWGQALELRKDGVRHCEGRVEYKHQGWGTVDGYRWTLKDASVVC 67  
 QY 88 QLCGPFAM----FRGQVTRHGIWLDVSCYNESALWECQ-----REWSHCYHGE 141  
 Db 68 QLCGCAAIAGFGGAYFGPL--GPIWLLYTSCEGTSTVSDCEHSNIDKRYNDGYNHGR 124  
 QY 142 DVGVCYGEANLRLVDGNNSCGRVEYKFOERWCTICDDGWNNTAAVVCROLGCPSS 201  
 Db 125 DAGVCSG-----FVRLAGDGPSCGRVEYHSGEAMIPVSDGNTLATAQIICAEIGCGKA 180  
 QY 202 FISSGVNVPAYLRPIWLDLILCOGNELALWNCRHGRGNHDCSHNEDYTLTCYDSSDLE 261  
 Db 181 VSVLGHLPRESSAQWAEERFCEGEPELWVCPVPCPGGTCHHSGSAQVVCYSEVR 240  
 QY 262 LRLVGTNRNCRVRELKIOGRNGTVCHHKWNAADVYCKQLGCGTALHFAGLPHLOSQ 321  
 Db 241 L-MTNGSSCOCEGOVEMNTSGORALCASHWSLANANVICROLGCGVATSTPGPHLVEEG 299  
 QY 322 DVVWLDGVSCNESPLDCRHSRTVNFCDLHNDVSVICSDGADLELRADGNNCSGR 381  
 Db 300 DQILTARFHCSEFLWSFVPTALGGPDCSHGNTASVICSN----- 342  
 QY 382 VEVRIHQWMTICDONKWEALVYCKQLGCPFSVFGSRRKPSNEARDIWNISICTGN 441  
 Db 343 -QIQVLQ-----CNDV-----SQPTGSAAS----- 363  
 QY 442 ESALWDCYDGAKRKTCPRRSRSDAGVTCSDADLRLVGAHSPCYGRLEVKYQGEWTV 501  
 Db 364 -----EDAPYCSDR--QLRLVDGGPCAGRVEILDQGSWGTIC 401  
 QY 502 HDRWSTRNAAVYCKQLGCGKPHFVGMTYFKESGPIWLDVSYICGNESNIWDCEHSG 561  
 Db 402 DGGWLDLDDARVVCRLGCGCEALNATGSAHFGAGSGPIWLDNLNCTCKESHVWRCPRGW 461  
 QY 562 KNCVHREDVYVTCSDATWGLRLVGGSNRCSGRLEVEYFQGRWGTVCDDGWSKAAVVC 621  
 Db 462 QHNCRHQDAGVTC--EFLALRVSEDOQCAGWLEVFYNGTWGVCYCRNPMEDITVSTIC 519  
 QY 622 SOLDCPSSIIGLGNAST---GYCKIWLDDVSCDGDDESILWSCRNSGWNDCSHSD 677  
 Db 520 ROLGCGDS-----GTLNNSVALREGFRPQWDRICQKKTDSLWQCPSPDPWYNNSCPREE 575  
 QY 678 VGVICSDADMEILRLVGGSSRCAGKVEVNVQAVGTILCANGMGNIAEVYRQLCEGSAI 737  
 Db 576 AXIWCADS--RQIRLVGGGRCGRVEILDQGSWGTICDDRWLDLDDARVVCQLGCGEAL 633  
 QY 738 RVSREPHEFTLHILMNSGCTGGEASLWDCIRWEKQTACHLNNEASLICSIAHRQRL 797

Db 634 DATVSSFFGTGSGPIWLDEVNCRGEESQVWRCPSGWRQHNCNHQEDAGVICSGF--VRL 691  
 QY 798 VCADPCSGRVEYKHADTWRSVCDSDFSLHAANVLCRELNCGDALSLSDHDEHFGNGLT 857  
 Db 692 AGGDGPGCSGRVEYHSGEAWTPVSDGNFTLPTAQVICAELGCGKAVSLVGLHMPFRESGGV 751  
 QY 858 WAEKQCSESETHALCPIVQHPEDTCHSREVGVVCSRYTDVRLV--NGKSQCDGQVEIN 916  
 Db 752 WAEERFCDGGEPELWSCPRVPCPGCTCLHSGAAQVVCVYTEVQLMKNKTSQCEGVEMK 811  
 QY 917 VLGHWSLCDTHWDPEDARVLCRQLSCGTALTSTTGKYGIGRSVRVWGHFRHCLGNESLL 976  
 Db 812 ISGRWRALCASHWSLANANVVCRLGCGVATSTPRGPHLVEGGDQISTAQFHCSEAESL 871  
 QY 977 DNCQMTVLGAPPCIHGNTSVICTGSLTQPLPCLANVSDPYLSAVPEGSALICLEDKRL 1036  
 Db 872 WSCPVTALGGPDCSHGNTASVICSGNHTQVLPQCNDFLSQPAGSRAESESPPYCSDSRL 931  
 QY 1037 RLVDGDSRCAGREYTHDGFNGTICDDGWDLSDAHVVCOKLGGCVAFNATVSAHFGEGSG 1096  
 Db 932 RLVDGGGCGGRVEILDQGSWGTICDDDDWLDLDARVVCRLGCGEALNATGSAHFGAGSG 991  
 QY 1097 PIWLDLNCNTGTESHLMQCPSGRWGQHDCHRKEDAGVICSEFTALRLYSETETESAGRL 1156  
 Db 992 PIWLDLNCNTGTESHLMQCPSGRWGQHDCHRKEDAGVICSEFTALRMVS--EDQOCAGWL 1049  
 QY 1157 EYFNGTWGSGVCRNITTAIAGIVCROLGCGGVVSLAPLSKTGSGFMWVDDIOCPKTH 1216  
 Db 1050 EYFNGTWGSGVCRNITTAIAGIVCROLGCGGVVSLAPLSKTGSGFMWVDDIOCPKTH 1109  
 QY 1217 ISIWOLCLAPWERRISSPAETWITCE-----DRIRVRGGDTESGRVE 1260  
 Db 1110 TSLWQCPSPGWKYSKSPKEAYISCEGRRPKSCTAACTOREKLRLRGSGSECSGRVE 1169  
 QY 1261 IWHAGSWGTVCDSDNDLAEAEVVCQQLGCGSALALRDASFQGGTGTIWLDMRCKNGNS 1320  
 Db 1170 VWHNGSWGTVCDSDSLAEAEVVCQQLGCGSALALRDASFQGGTGTIWLDMRCKNGNS 1229  
 QY 1321 FLWDCHAKPWGSDCHGKEDAGVRCGSG-----QSLKSLNASSGHLLALIL 1364  
 Db 1230 SLWDCAEPWGSQDCKHEEDAGVRCGSGVTTPTTAGTRTTSNLSLPGTSPGLCLIL 1289  
 QY 1365 SSIFGLLLVLFLVLTWCVRQKHPLRLVSTRRGSLEENLFHEMETCLRKREDPHGTR 1424  
 Db 1290 GSLLFLVLVLTQLLRW--RAER-----RALSSVEDALAEAYEBDYLLTQKEGLGSP 1342  
 QY 1425 TS-DDTPNHGCEADSTSLGLVLPASEATK 1453  
 Db 1343 DQMTDVPDENYDDAEVVPVTPSPSQGNE 1372  
 RESULT 2  
 NETR\_HUMAN  
 ID NETR\_HUMAN STANDARD; PRT; 875 AA.  
 AC P56730; Q90U16;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Neurotrophin precursor (EC 3.4.21.-) (Motopsin) (Leydin).  
 GN PRSS12.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98201705; PubMed=9540828;  
 RA Proba K., Gschwend T.P., Sonderegger P.;  
 RT "Cloning and sequencing of the cDNA encoding human neurotrophin.";  
 RL Biochim. Biophys. Acta 1396:143-147(1998).  
 RN [2]



RC TISSUE=Brain;  
 RA MEDLINE=97401523; PubMed=9245503;  
 RA Gschwend T.P., Krueger S.R., Kozlov S.V., Wolfer D.P., Sonderegger P.,  
 RT "Neurotrophin, a novel multidomain serine protease expressed in the  
 RT nervous system.",  
 RL Mol. Cell. Neurosci. 9:207-219(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98008848; PubMed=9344839;  
 RX Yamamura Y., Yamashiro K., Tsuruoka N., Nakazato H., Tsujimura A.,  
 RA Yamaguchi N.;  
 RT "Molecular cloning of a novel brain-specific serine protease with a  
 RT kringle-like structure and three scavenger receptor cysteine-rich  
 RT motifs.",  
 RL Biochem. Biophys. Res. Commun. 239:386-392(1997).  
 CC -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC  
 CC ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH  
 CC LEARNING AND MEMORY OPERATIONS.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN CEREBRAL CORTEX, HIPPOCAMPUS  
 CC AND AMYGDALA.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 3 SRCR DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y13192; CAA73646.1; .  
 DR EMBL; D89871; BAA23986.1; .  
 DR HSP; P00763; IDPO.  
 DR MEROPS; S01.237; .  
 DR MGD; MGI:1100881; Prss12.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Ser\_protease\_Try.  
 DR Pfam; PF0001190; Srcr\_receptor.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR Pfam; PF00530; SRCR; 3.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE; FALSE\_NEG.  
 DR PRINTS; PR00258; SPERACTRCPT.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00202; SR; 3.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; FALSE\_NEG.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 DR PROSITE; PS00420; SRCR\_1; 3.  
 DR PROSITE; PS0287; SRCR\_2; 3.  
 DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 761  
 FT DOMAIN 85 157  
 FT KRINGLE.  
 FT DOMAIN 166 267  
 FT SRCR 1.  
 FT DOMAIN 273 373  
 FT SRCR 2.  
 FT DOMAIN 386 487  
 FT SRCR 3.  
 FT DOMAIN 505 761  
 FT SRCR PROTEASE.  
 FT DOMAIN 505 516  
 FT ZINCEN ACTIVATION REGION.  
 FT ACT\_SITE 516 517  
 FT REACTIVE BOND (POTENTIAL).  
 FT ACT\_SITE 562 562  
 FT CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 612 612  
 FT CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 711 711  
 FT POTENTIAL.  
 FT DISULFID 505 636

FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 521 521 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 761 AA; 84118 MW; DF507B03712164E6 CRC64;  
 Query Match 8.8%; Score 719.5; DB 1; Length 761;  
 Best Local Similarity 31.8%; Pred. No. 4.8e-43;  
 Matches 191; Conservative 80; Mismatches 233; Indels 97; Gaps 19;  
 QY 451 DCKAKRTCFRRSDAGVI-----CSDKADLD-LRLVGAHSPCYGRLEVKYQGEWGTVCVD 503  
 DB 134 DSGRPWCFCYRNAQCKVDWGYDCGCGPALPVIRLVGGNSGHEGRVELYHAGOWGTICDD 193  
 QY 504 RWSTNAAVVCKLGGKPMHVFMTYFKEASGPTLWDDVSCIGNESNTWDCHEHGWGKH 563  
 DB 194 QWMDADADVICRLGLSIGIAKAWHAHFGSGSPILLDEVRCSTNELSTEQCPKSWGSH 253  
 QY 564 NCVHREDVIVTCSGDATWGLRLVGGSNCRSGRLEVYFGFRMGTVGDDGWSKAAAVVCSQ 623  
 DB 254 NCGHEDAGVSCVPLTDGVIRLAGGKSTHEGRLEVYFGQWGTVCDDGTENTVYVAC-- 311  
 QY 624 LDCPSIIIGMLGNAST-----GYGK-IWLDVDCDGDSDLSWSCRNSGWNDCSHSD 677  
 DB 312 -----RLLGFKYKQSSVNHFDGSRNPWLDDVSCGREVSFIQCSRRQWRHDCSHRD 366  
 QY 678 VGVICSDASD-----NELRVGGSSRCAGKVEVNVGAVGILCANGWGMIAEVCVRQ 730  
 DB 367 VGLTCYPDSGHRSLSPGFPPIRLVDGENKEGRVEVFVQWGTICDDGWTDKHAAVICRQ 426  
 QY 731 LECGSAIRVSRPHPTERTLHILMSNGCTGGEASLWDCIRWEWKOTACHLNNEASLICS 790  
 DB 427 LGYKPARARTWAYGEGKPIHMDNVKCTGNEKALDCVKODIGHNCRHSEDAGVICD 486  
 QY 791 -----AHRQRLVGDMPGCSG-----RVEYKHADTWRS 818  
 DB 487 YLEKKAASSGNKEMLSGCGRLRLHRRQKRIITGGNSLRGAWPQASLRSAHGD-GRL 545  
 QY 819 VC-----DSDFSLHAANVLCRELNCDAISLSVGDH-----FGKNG--LTWAEKQ 863  
 DB 546 LCGATLLSSCWVLTAAHCFKRYGNSRSYAVRVGYDHTLVPEEFQEIQQVQIVTHRNTR 605  
 QY 864 CSGSETHALCPVQHPEPTCIH-SREVGVCVSRYTDLVRLVNGKSCQDCQVENVLGHWG 922  
 DB 606 PRSDYDIALVRL-QGPGQCARLSTHVLPACLPLWRERPKQTASNC-----HIIG-WG 657  
 QY 923 SLCDTHWDPEDARVLCRLQSLCTALSTGCKYIGERSVVRWGHFRHCLGN-----ESLLDNC 979  
 DB 658 -----DTGRAYSRITQ-QAAVPLLPKFCRKERYKGLFTGRMLCAGNLOEDNRVDC 707  
 QY 980 Q 980  
 DB 708 Q 708  
 RESULT 4  
 CD5L\_HUMAN  
 ID CD5L\_HUMAN STANDARD; PRT; 347 AA.  
 AC 043866;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE CD5 antigen-like precursor (Sp-alpha) (CT-2) (Igm-associated peptide).  
 GN CD5L.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF N-TERMINUS, AND TISSUE SPECIFICITY.  
 RC TISSUE=Spleen;  
 RX MEDLINE=9719777; PubMed=9045627;  
 RA Gebe J.A., Kiener P.A., Ring H.Z., Li X., Francke U., Aruffo A.;  
 RT "Molecular cloning, mapping to human chromosome 1 q21-q23, and cell



```

DR EMBL; AF011428; AAD01445.1; -.
DR EMBL; BC006799; AAH06799.1; -.
DR MGD; MGI:1333862; Ap16.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00530; SRCR; 3.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR; 3.
DR PROSITE; PS00420; SRCR.1; 1.
DR PROSITE; PS0287; SRCR.2; 3.
KW Signal; Repeat; Glycoprotein; Apoptosis; Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 352
FT DOMAIN 27 128
FT DOMAIN 141 241
FT DOMAIN 246 348
FT CARBOHYD 99 99
FT CARBOHYD 195 195
FT CARBOHYD 229 229
FT CARBOHYD 316 316
FT VARIANT 61 61
FT VARIANT 197 197
FT VARIANT 205 205
FT CONFLICT 13 13
FT CONFLICT 113 113
FT CONFLICT 352 AA; 38835 MW; 414984ACEFI3B23 CRC64;
SQ SEQUENCE 352 AA; 38835 MW; 414984ACEFI3B23 CRC64;

Query Match 8.2%; Score 664.5; DB 1; Length 352;
Best Local Similarity 39.1%; Pred. No. 1.5e-39;
Matches 131; Conservative 55; Mismatches 134; Indels 15; Gaps 9;

QY 253 TCYSDLELRVLTGNTNRCMVEKIQGRGTCTVCHHKWNAADVCKQCCGTALHF- 311
DB 18-SCFSESPTRVQLVGAHRCRGEVVEHNGQWGTCDGWDNRDVAVVCRELNCGAVIQTP 77

QY 312 AGLPHLQSGSDV-VMLDGVSCSGNESFLWDCRHSFTVNFCDLHONDVSVICSD-GADL-- 367
DB 78 RGASYQPPASEQRVLQGGVDCNGTETLAQCELNDV-FDCSHEEDAGACQENPDSDLF 136

QY 368 ---ELRADGSNNCSGRVREIHEQWTTICDQNWKNQALVYVCKQCGPFSVFGSRRAP 424
DB 137 IPEDVLVDGPGHCQGRVLEVHQSQWSTVCAGWNLQVSKVCRQLGCGRALLTYGSCNK 196

QY 425 SNEAR-DIWINSTGNSALWDCTYDGKAKTFCRRSDAGVICSADKADLLDLVGAHS 483
DB 197 STQKGPIWMGKMSQGANLRSL-SRLNNCTHGEDTWMECD--PFELKLVGGDT 253

QY 484 PCVGRLEVYQSGWGTVCVCHDRSTRNAAVVCKQCGKPMHVFGMT--YFKEASGPFIWD 541
DB 254 PCSGRLEVHLKGSWGSVCDDNNGEKEDQVYVCKQCGKSLHPSPKTKIYGPAGRIWLD 313

QY 542 DVCSCIGNESNWDCEHSGWKKHNCVHREDVIVTCS 576
DB 314 DVNCSGKEQSLEFCRHLRGYHDTCKEDVEVICT 348

RESULT 6
LOL4_MOUSE STANDARD; PRT; 757 AA.
AC 924C6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lysyl oxidase homolog 4 precursor (EC 1.4.3.-) (Lysyl oxidase-like protein 4) (Lysyl oxidase related protein C).
GN LOL4 OR LOXC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=21316447; PubMed=11292829;
RX ITO H., Akiyama H., Iguchi H., Iyama K., Miyamoto M., Ohsawa K.,
RA ITO H., Akiyama H., Iguchi H., Iyama K., Miyamoto M., Ohsawa K.,

```

```

RA Nakamura T.;
RT "Molecular cloning and biological activity of a novel lysyl oxidase-
RT related gene expressed in cartilage.";
RL J. Biol. Chem. 276:24023-24029(2001).
CC -I- FUNCTION: May modulate the formation of a collagenous
CC extracellular matrix.
CC -I- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTO) (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Extracellular (Potential).
CC -I- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
CC A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
CC LYSINE (BY SIMILARITY).
CC -I- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC -I- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF338440; AAK71933.1; -.
CC MGD; MGI:1914823; Lox14.
CC InterPro; IPR001695; Lysyl_oxidase.
CC InterPro; IPR001190; Srcr_receptor.
CC Pfam; PF01186; Lysyl_oxidase; 1.
CC Pfam; PF00530; SRCR; 4.
CC PROSITE; PS01387; Lysyl_oxidase; 1.
CC PROSITE; PS00926; LYSYL_OXIDASE; FALSE_NEG.
CC PROSITE; PS00420; SRCR.1; 1.
CC PROSITE; PS0287; SRCR.2; 4.
KW Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 25
FT CHAIN 26 757
FT DOMAIN 33 134
FT DOMAIN 160 288
FT DOMAIN 312 412
FT DOMAIN 422 530
FT DOMAIN 534 737
FT METAL 612 612
FT METAL 614 614
FT METAL 616 616
FT MOD_RES 639 639
FT MOD_RES 675 675
FT CARBOHYD 199 199
FT CARBOHYD 615 615
FT CARBOHYD 630 630
SQ SEQUENCE 757 AA; 84705 MW; D9861368F63B7B2A CRC64;

Query Match 7.8%; Score 633.5; DB 1; Length 757;
Best Local Similarity 32.0%; Pred. No. 5.2e-37;
Matches 165; Conservative 67; Mismatches 190; Indels 93; Gaps 17;

QY 42 NGTDLRLVN-GDGPCTGVVEKFGQGWGTCDGWNNTASTVYVCKQCGPFSFAM--- 97
DB 28 SGTK-KLRVGPADRPPEERLEVLHQGGTVCDDDFALQATVACRQLG--FESALTA 84

QY 98 --FRFGQAVTRHGKIWLDDVSCYGNESALWECQHRWESHNCYHGVDGVNCGEALGL 155
DB 85 HSAKYQGQ---EGPIWLDNVRCLGTEKTLDCQSGNMGWISDCRHSDEGVVCHPRQHG 141

QY 156 RLVDGNNSS-----CSGRVEVKFQERWGTICDGNLNTA 189
DB 142 HSEKVSNALPGQRRLEEVRLKPIILASAKRHSPVTEGAVEVRYDGHWRQVCCQWTMNS 201

QY 190 AVVCROLGCPSS-SFTSS-----GVNNSPAVLRPIWLDDILCOGNELALWN 233
DB 202 RVVCGMLGFPSTQSVNSHYRKVWNLKMKDPKSRNLSTKKNKSFVHRVDFGTEPHLAK 261

QY 234 CR-----HRGWGNHDCSHNEDVTLTCY-----DSSDLELRVLTGNTNRC 271

```

Db 262 CQVAVPGRGKLRPACPGHIAVVCVAGPHFRQKPKPTKESHAEELKVLRLSGAQVG 321  
Qy 272 MGRVELKIQGRWGTCHHKWNNAADVCKQLGCGTALHFAGLPHLQSGDVVWLDGVSC 331  
Db 322 EGRVFLNMQWGTCDHRNLISASVVCQLGFGSREALFGAQLCGGLPIHLSEVRC 381  
-Qy 332 SGNESFLWDC-RHSGTGVNFDCLHNDVSVICS---DGADLELRADSGNCSGRVEVRIH 387  
Db 382 RGYERTLGDCLALGSON-CQGHANDAARCNIPDMGFQNKVRLAGGRNSEEGVVEQVE 440  
Qy 388 ----EQHWTTICDQWKNNEQALVVCCKQLGCPFSVFGSRA---KPSNEARDIWNISICTG 440  
Db 441 VNGGPRNGTVDGKAKRTCFR---RSDAGVICSDKA 472  
Qy 441 NESALWDCCTVDGKAKRTCFR---RSDAGVICSDKA 472  
Db 501 TEMALQOCQRHGPVH---CSHGPGRFSAGVACMNSA 533

RESULT 7  
LOL3\_MOUSE STANDARD; PRT; 754 AA.  
AC Q22175; Q9JJ39;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lysyl oxidase homolog 3 precursor (EC 1.4.3.-) (Lysyl oxidase-like protein 3) (lysyl oxidase related protein 2).  
GN LOXL3 OR LOXL2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J, and 129/SvJ; TISSUE=Muscle;  
RX MEDLINE=99126643; PubMed=9927484;  
RA Jang W., Hua A., Spilson S.V., Miller W., Roe B.A., Meisler M.H.;  
RT "Comparative sequence of human and mouse BAC clones from the mmd2 region of chromosome 2p13.";  
RL Genome Res. 9:53-61(1999).  
CC -1- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTQ) (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Extracellular (Potential).  
CC -1- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF A TYROSINE RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A LYSINE (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 4 SRCR DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: AF0533368; AAC83205.1; -  
EMBL: AF084363; AAC95338.1; -  
DR MGD; MGI:1337004; Loxl3.  
DR InterPro; IPR001695; Lysyl\_oxidase.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00530; SRCR; 4.  
DR Pfam; PF01186; Lysyl\_oxidase; 1.  
DR PRINTS; PR00074; LYSYLOXIDASE.  
DR PRINTS; PR00258; SPERACTRCPTR.  
DR ProDom; PD013887; Lysyl\_oxidase; 1.  
DR SMART; SM00202; SR; 4.  
DR PROSITE; PS00926; LYSYL\_OXIDASE; 1.  
DR PROSITE; PS00420; SRCR\_1; 1.  
DR PROSITE; PS0287; SRCR\_2; 4.  
KW Oxidoreductase; copper; Glycoprotein; Repeat; Signal.  
FT SIGNAL 1 26 POTENTIAL.

FT CHAIN 27 754 LYSYL OXIDASE HOMOLOG 3.  
FT DOMAIN 45 146 SRCR 1.  
FT DOMAIN 170 283 SRCR 2.  
FT DOMAIN 308 408 SRCR 3.  
FT DOMAIN 418 526 SRCR 4.  
FT DOMAIN 530 733 LYSYL-OXIDASE LIKE.  
FT METAL 608 608 COPPER (POTENTIAL).  
FT METAL 610 610 COPPER (POTENTIAL).  
FT METAL 612 612 COPPER (POTENTIAL).  
FT MOD\_RES 635 635 CROSS-LINKED TO TYROSYLQUINONE (BY SIMILARITY).  
FT MOD\_RES 671 671 TYROSYLQUINONE (BY SIMILARITY).  
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 626 626 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 754 AA; 83681 MW; F8758400943F4FF2 CRC64;

Query Match 7.7%; Score 629; DB 1; Length 754;  
Best Local Similarity 31.0%; Pred. No. 1.1e-36;  
Matches 155; Conservative 72; Mismatches 183; Indels 90; Gaps 16;

Qy 56 FCSGTVEYKFOGQGTVCDDGWNTTASTVCKQLGCPFSFA-----MRFQOAVTRHGKI 110  
Db 54 FYEGRVEIQRAGEWGTICDDFTLQAAHVLCRELG--FTEATGWTSHAKYGPQT---GRI 108  
Qy 111 WLDVSCYGNESALWECQHWGSHNCHYHGEDVGVNC-----YGEANL----- 153  
Db 109 WLDNLSRGTGSGYTECASRGWNSDCTHDEDAGVICKDRLPGFSDSNVIEVHQLOVE 168  
Qy 154 -----GLRLVDGNNSCSGRVEYKFOERWGTICDDGNLNTAAVVCQLGCPSS-- 201  
Db 169 EVRLRPAVEWGRRLP---VTEGLVEVRLPGWQVCDKGSANSHSVVVCMLGFPGEKR 225  
Qy 202 -----FISGGVYNPAVLRIWLDLILCOGNELALMNCRRHGWGHHND---CSHNEDVT 251  
Db 226 VNMAFYRLAQKQHS-----FGLHSVACVQTEAHLSCLSLEFYRANDTTTRCSGPNPAV 279  
Qy 252 LTCY-----DSSDLELRVGGTNCRCMGRVELKIQGRWGTCHHKWN 293  
Db 280 VSCVLGPLYAFYFTGOKKQKHPOGEARVRLKGAGHOGEGRVEVLKAGTGTVCDDKWL 339  
Qy 294 AAADVCKQLGCGTALHFAGLPHLQSGDVVWLDGVSCGNESFLWDCRHSCTGVNFDCLH 353  
Db 340 QAASVVCPELGFGTAREALSARGMQCGNGAIHLSEVRCSGOEPSLWRCPSKNITAECSH 399  
Qy 354 QNDVSVICS---DGADLELRADSGNCSGRVEVRI-----HEQWTTICDQWKNNEQALV 406  
Db 400 SQDAGVRCNLPTGYETKIRLSGGRSRYEGRVEVQIGIPGHLRMGLICGDDMGTLKAV 459  
Qy 407 CKQLGCPFSVFGSRA---KPSNEARDIWNISICTGNESALWDCYTDGKAKRTCFR---R 461  
Db 460 CRQLGLGVANHLGLETWYDWSGVNTEVVMVSGVRCITGSELSNQCAHH--SSHITCKKTGR 518  
Qy 462 SDAGVICSKADLRLVGA 481  
Db 519 FTAGVICSETAS-DLLHSA 537

RESULT 8  
LOL3\_HUMAN STANDARD; PRT; 753 AA.  
AC P58215; Q96RSL;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lysyl oxidase homolog 3 precursor (EC 1.4.3.-) (Lysyl oxidase-like protein 3).  
GN LOXL3 OR LOXL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21181571; PubMed=11284725;  
 RA Maki J.M., Kivirikko K.I.;  
 RT "Cloning and characterization of a fourth human lysyl oxidase  
 isoenzyme.";  
 RL Biochem. J. 355:381-387(2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21280915; PubMed=11386757;  
 RA Jourdan-Le Saux C., Tomshe A., Ufalusi A., Jia L., Csizsar K.;  
 RT "Central nervous system, uterus, heart, and leukocyte expression of  
 the LOXL3 gene, encoding a novel lysyl oxidase-like protein.";  
 RL Genomics 74:211-218(2001).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21233589; PubMed=11334717;  
 RA Huang Y., Dai J., Tang R., Zhao W., Zhou Z., Wang W., Ying K., Xie Y.,  
 MAO Y.;  
 RT "Cloning and characterization of a human lysyl oxidase-like 3 gene  
 (hLOXL3).";  
 RL Matrix Biol. 20:153-157(2001).  
 CC -1- COPACITOR: COPPER AND LYSINE TYROSYLQUINONE (LTO) (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Extracellular (Potential).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES, THE HIGHEST LEVELS  
 AMONG THE TISSUES STUDIED BEING IN THE PLACENTA, HEART,  
 OVARY, TESTIS, SMALL INTESTINE AND SPLEEN.  
 CC -1- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF  
 A TYROSINE RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A  
 LYSINE (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 4 SRCR DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF282619; AAK51671.1; -;  
 DR EMBL; AF311313; AAK63205.1; -;  
 DR EMBL; AF284815; AAK91134.1; -;  
 DR Genew; HGNC:13869; LOXL3.  
 DR InterPro; IPR001695; Lysyl\_oxidase.  
 DR InterPro; IPR001190; Srcr\_receptor.  
 DR Pfam; PF00530; SRCR; 4.  
 DR Pfam; PF01186; Lysyl\_oxidase; 1.  
 DR PRINTS; PR00074; LYSILOXIDASE.  
 DR PRINTS; PR00258; SPERACTRCPTR.  
 DR ProDom; PD013887; Lysyl\_oxidase; 1.  
 DR SMART; SM00202; SR; 4.  
 DR PROSITE; PS00926; LYSYL\_OXIDASE; 1.  
 DR PROSITE; PS00420; SRCR\_1; 1.  
 DR PROSITE; PS0287; SRCR\_2; 4.  
 KW Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 26 753  
 FT DOMAIN 44 145  
 FT DOMAIN 169 282  
 FT DOMAIN 307 407  
 FT DOMAIN 417 525  
 FT DOMAIN 529 732  
 FT METAL 607 607  
 FT METAL 609 609  
 FT METAL 611 611  
 FT MOD\_RES 634 634  
 FT MOD\_RES 670 670  
 FT CARBOHYD 111 111  
 FT CARBOHYD 266 266

FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 159 159 E -> K (IN REF. 3).  
 SQ SEQUENCE 753 AA; 83166 MW; 582C46DA25E05A69 CRC64;  
 Query Match 7.68; Score 622.5; DB 1; Length 753;  
 Best Local Similarity 31.28; Pred. No. 3.1e-36;  
 Matches 167; Conservative 71; Mismatches 202; Indels 95; Gaps 19;  
 QY 30 LLLNSCFLLISFNFGT-----DLELRVN-GDGPCSGTVEVKFGQWGTVD 74  
 DB 14 LLL--CLLCSCSLGSPSTGPEKAGSQGLFRFLAGEPRKPYEGRVIQIRAGETICD 71  
 QY 75 DGWNTASTVVKVQKLGCPFSFA-----MPREGQAVTRHGKIWLDDVSCYNESALWECOH 129  
 DB 72 DDFTLQAAHILCRELG--FTEATGWTHTSAKYGPGT--GRIWLDNLSCGTSEQSTECAS 126  
 QY 130 REMGSHNCHGEDVGVNC-----YGEANL-----GLRLVDGNN 162  
 DB 127 KGWNSDCTHDEDAGVICKDORLPGSDSNVIEVEHHLOVEVRIRPAVGWGRRLP 183  
 QY 163 SCSGRVEYKFORWGTICDDGNLNTAAVVCRLQPCSPSSFISSGVVNSPAVL 219  
 DB 184 VTEGLVEVRLPDGWSQVCDKGSASHNVHVCMLGFPSEKRVNAAFYRLAQROHSEGL 243  
 QY 220 DDLCOGNELALWNRHGRGNHD---CSHNEDEVLTLC-----YDSS----- 258  
 DB 244 HGVACVGTEAHLSLGSEFYRANDTARCPGGPVPVPGPVYAASSGQKQKQSKPG 303  
 QY 259 DLELRVGVTRCMGRVELKIOGRWGTVCHHKWNAADVVCOLGCGTALHAGLPHLQ 318  
 DB 304 EARVRLKGAHPGEGRVEVLKASTWGTVCDRKWDLAHVASVVCRELGFSARALSGRMG 363  
 QY 319 SGSDVYVLDVSCSGNESFLWDCRHSGTVNFDCLHQNVDVICS---DGADLELRADGS 375  
 DB 364 QGMGAILHSEVRCSEQLSLNCKPHKNITAEPCSHSQDAGVRCNLPTGTAEIRLSGGR 423  
 QY 376 NNCSSRVEVRIHE-----QWTTICDQNKNEQALVCKQCPFSVFGSRA--KFSNEAR 429  
 DB 424 SQHEGRVEYQIGPGPLRMGLGICDDWGTLEAMVACRQLGLGYANGHGLQETWYDWSGNIT 483  
 QY 430 DIWINSISCTGNESALWDCTYDGKAKRTCFR---RSDAGVICSADLRLVGA 481  
 DB 484 EVVMGVRCTGTGTELSLDOCAHHG-THICTKRTGTRTAGVICSETAS-DLLHSA 536  
 RESULT 9  
 ID LOLA\_HUMAN STANDARD; PRT; 756 AA.  
 AC Q96JB6; Q96PC0; Q96DY1;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lysyl oxidase homolog 4 precursor (EC 1.4.3.-) (Lysyl oxidase-like  
 protein 4) (Lysyl oxidase related protein C).  
 GN LOXL4 OR LOXC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Testis;  
 RX MEDLINE=21316447; PubMed=11292829;  
 RA Ito H., Akiyama H., Iguchi H., Iyama K., Miyamoto M., Ohsawa K.,  
 Nakamura T.;  
 RT "Molecular cloning and biological activity of a novel lysyl oxidase-  
 related gene expressed in cartilage.";  
 RL J. Biol. Chem. 276:24023-24029(2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21550107; PubMed=11691589;

RA Maki J.M., Tikkanen H., Kivirikko K.I.;  
 RT "Cloning and characterization of a fifth human lysyl oxidase  
 RT isoenzyme: the third member of the lysyl oxidase-related subfamily  
 RT with four scavenger receptor cysteine-rich domains.";  
 RL Matrix Biol. 20:493-496(2001).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Placenta;  
 RC TISSUE=Eye;  
 RA Asuncion L.P., Fogelgren B., Fong K.S.K., Fong S.F.T., Kim Y.,  
 RA Csizsar K.;  
 RT "A novel human lysyl oxidase-like gene (LOXL4) on chromosome 10q24 has  
 RT an altered SRCR domain and is down-regulated by H-ras.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Eye;  
 RC TISSUE=Eye;  
 RA Strausberg R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -|- FUNCTION: May modulate the formation of a collagenous  
 CC extracellular matrix.  
 CC -|- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTO) (BY SIMILARITY).  
 CC -|- SUBCELLULAR LOCATION: Extracellular (Potential).  
 CC -|- TISSUE SPECIFICITY: Expressed in many tissues, the highest levels  
 CC among the tissues studied being in the skeletal muscle, testis and  
 CC pancreas. Expressed in cartilage.  
 CC -|- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF  
 CC A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A  
 CC LYSINE (BY SIMILARITY).  
 CC -|- SIMILARITY: CONTAINS 4 SRCR DOMAINS.  
 CC -|- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF38441; AAK71934.1; -  
 CC EMBL: AY036093; AAK64186.1; -  
 CC EMBL: AF395336; AAL27543.1; -  
 CC EMBL: BC013153; AAH13153.1; -  
 CC Genew: HGNC:17171; LOXL4.  
 CC InterPro: IPR001695; Lysyl\_oxidase.  
 CC InterPro: IPR001190; Srcr\_receptor.  
 CC Pfam: PF01186; Lysyl\_oxidase; 1.  
 CC Pfam: PF00530; SRCR; 4.  
 CC ProDom: PD013887; Lysyl\_oxidase; 1.  
 CC PROSITE: PS00926; LYSYL\_OXIDASE; FALSE\_NEG.  
 CC PROSITE: PS00420; SRCR.1; 1.  
 CC PROSITE: PS50287; SRCR.2; 4.  
 CC Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.  
 KW SIGNAL  
 FT 1 24 POTENTIAL.  
 FT CHAIN 25 756 LYSYL OXIDASE HOMOLOG 4.  
 FT DOMAIN 32 133 SRCR 1.  
 FT DOMAIN 159 287 SRCR 2.  
 FT DOMAIN 311 411 SRCR 3.  
 FT DOMAIN 421 529 SRCR 4.  
 FT DOMAIN 533 736 LYSYL-OXIDASE LIKE.  
 FT METAL 611 611 COPPER (POTENTIAL).  
 FT METAL 613 613 COPPER (POTENTIAL).  
 FT METAL 615 615 COPPER (POTENTIAL).  
 FT MOD\_RES 638 638 CROSS-LINKED TO TYROSYLQUINONE (BY  
 FT SIMILARITY).  
 FT MOD\_RES 674 674 TYROSYLQUINONE (BY SIMILARITY).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 3 3 W -> R (IN REF. 4).  
 FT CONFLICT 101 101 R -> Q (IN REF. 4).  
 FT CONFLICT 405 405 D -> A (IN REF. 4).  
 FT CONFLICT 493 493 S -> G (IN REF. 3).  
 FT CONFLICT 539 539 A -> T (IN REF. 3).

FT CONFLICT 542 542 V -> A (IN REF. 3).  
 FT CONFLICT 703 703 Y -> H (IN REF. 3).  
 SQ SEQUENCE 756 AA; 84483 MW; 13051ACADB922BBC CRC64;  
 Query Match 7.6%; Score 618; DB 1; Length 756;  
 Best local similarity 31.2%; Pred. No. 6.4e-36;  
 Matches 159; Conservative 68; Mismatches 191; Indels 92; Gaps 16;  
 QY 47 ELRLVNGDG-PCSGTVEVKFQGGMTVCDDGNTTASTVVCVKQLGCPFSFAM-----FRF 100  
 Db 31 KLRLLPESPEGRLEVLHVGQGTVCDDNFALQEAQVACRQLG--FEALTWHAHSAKY 88  
 QY 101 GOAVTRGKIWLDDVSCYGNESALWECQHRWGNHNCYHGEGDVGVNCYGEANGLRLVDG 160  
 Db 89 GQG---EGPIWLDNVRVCVTGTESSLDQCGSNGWVSDCSHSEDVGVICHPRRHRYLSETV 145  
 QY 161 NNS-----CSGRVEVKEFQERWGTICDDGNLNTAAVVCV 194  
 Db 146 SNALGPQGRRLLEEVRLKPIILASAKQHSPTVEGAVKRYEGHWRQVCDQGTMMNSRYVCG 205  
 QY 195 QLGCPSSF-ISS-----GVNNSPAVLRLPIWLDLILCOGNELALWNCR--- 235  
 Db 206 MLGFPSEVPVDVSHYRKYRWDLKMRDPKSLKSLTNKNSFWIHOVTCUGTPEHMANCOVQV 265  
 QY 236 --HRGNHNDCHSHNEDVTLTCYD-----SSDLELRVLGGTNNRMRGE 276  
 Db 266 APARKLRPACPGMHAVVSCVAGPHPRPKTPQRKGSWAEPRVRLRSGAQVGEGRVE 325  
 QY 277 LKIQGRGVTCHHKWNNAAADVCKQLCGGTALHFAGLPHLQSGDVVWLDGVSCSNGES 336  
 Db 326 VLNRQMGVTCVDRHNLISAVVCRQLGFGSAREALFGARLQGLGPIHLSEVRCRGYER 385  
 QY 337 FLWDC-RHSQTVNFCLQNDVSVIC---SDGADLELRADGNSNCSGRVEVRTH----E 388  
 Db 386 TLDCEPALESQN-GCOHENDAAVRNVNPNWGFONQVRLAGRIPEGLLEQVGVNGVP 444  
 QY 389 QWMTICDQNNKNEQALVYCKQLGCPFSVFGSRRRA---KPSNEARDIWNISCTGNSAL 445  
 Db 445 RWGSCVSENKTEAMVACVACVGLGLGFAIHAYKETWMSGTPRAQVYVWVSGVRCSTELAL 504  
 QY 446 WDCTYDGAARTCFER---RSDAGVICSDKA 472  
 Db 505 QQCQRHGPVH--CSHGGRFLAGVSCMDSA 532  
 RESULT 10  
 CD6\_MOUSE  
 ID CD6\_MOUSE STANDARD; PRT; 665 AA.  
 AC Q61003; Q61004; Q60679;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE T-cell differentiation antigen CD6 precursor.  
 GN CD6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Thymocytes;  
 RX MEDLINE=96062291; PubMed=7594475;  
 RA Robinson W.H., Prohaska S.S., Santoro J.C., Robinson H.L.,  
 RA Parnes J.R.;  
 RT "Identification of a mouse protein homologous to the human CD6 T cell  
 RT surface protein and sequence of the corresponding cDNA.";  
 RL J. Immunol. 155:4739-4748(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RX MEDLINE=95174761; PubMed=7870060;  
 RA Whitney G., Bowen M., Neubauer M., Aruffo A.;  
 RT "Cloning and characterization of murine CD6.";  
 RL Mol. Immunol. 32:89-92(1995).





```

or send an email to license@isb-sib.ch).
-----
CC EMBL; X60992; CAA43306.1; -
DR EMBL; U66142; AAC51161.1; -
DR EMBL; U66143; AAC51162.1; -
DR EMBL; U66144; AAC51163.1; -
DR EMBL; U66145; AAC51164.1; -
DR EMBL; U66146; AAC51165.1; -
DR PIR; S26741; S26741.
DR MIM; HGNC:1691; CD6.
DR InterPro; IPR001190; Ssrcr_receptor.
DR Pfam; PF00530; SRCR_3.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR_3.
DR PROSITE; PS00420; SRCR_1; 1.
DR PROSITE; PS50287; SRCR_2; 3.
KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
KW Alternative splicing.
FT SIGNAL 1 17
FT CHAIN 18 668
FT DOMAIN 18 402
FT TRANSMEM 403 423
FT DOMAIN 424 668
FT DOMAIN 45 156
FT DOMAIN 161 260
FT DOMAIN 265 361
FT CARBOHYD 28 49
FT CARBOHYD 49 49
FT CARBOHYD 112 112
FT CARBOHYD 118 118
FT CARBOHYD 229 229
FT CARBOHYD 339 339
FT CARBOHYD 345 345
FT CARBOHYD 368 368
FT VARSPPLIC 431 462
FT VARSPPLIC 463 504
FT N -> D (IN ISOFORM CD6C AND ISOFORM
VFMPIQVAPPEDSDSDSDYEHYDFSAOPVPLTFY
CD6E).
FT MISSING (IN ISOFORM CD6D AND ISOFORM
CD6E).
FT MISSING (IN ISOFORM CD6D AND ISOFORM
CD6E).
FT VMLFT -> GPGPAP (IN REF. 1).
FT MISSING (IN REF. 2; AAC51162).
FT SEQUENCE 668 AA; 71828 MW; EFB434608012C49E CRC64;

Query Match
Best Local Similarity 28.7%; Score 592; DB 1; Length 668;
Matches 154; Conservative 59; Mismatches 140; Indels 184; Gaps 15;

QY 900 VRLVNGKSCDQGVNVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGKYGERS 959
DB 45 VRLTNGSSSCSGTVEVRLEASWEPACGALWDSRAEAVCRALCGGAEASQ----- 96
QY 960 VRVWGRHFCL-----GNESLLDNCOMTLVGAAPCIHGNVTYSVICTGSLTOP 1006
DB 97 -----LAPPTPELPPPPAAGNTSVAAAN--FLAGAP-----ALLCSGAENR- 135
QY 1007 LFPCLANVSDPYLSAYPEGSALICLEDKRLRLVDGSRGACGRVEIYHDFGFWETICDDGWD 1066
DB 136 -----LCEVVEHACRSRGRRARVTCANRLRLVDGGACAGRVEMLEHGEWSVCDDTWD 191
QY 1067 LSDAHVVCQKLGCGVAFNATVSAHFEGSGPIWLDLNTGTESHUWQCPSRGWGHDCR 1126
DB 192 LEDARVVCRLGCGWVAQVLPGLHFTPGRPIHRDQVNGSGAEAYLWDCPGLP-GQHYCG 250
QY 1127 HKDAGVICSEFTALRLYSETETESACRLEVEYNGTWGSGVRNITTAGIVCRLQGC 1186
DB 251 HKDAGVVCSE----- 261
QY 1187 GENGWVSLAPLSKTSFGFMVDDIQCPKTHISIWQCLSAFWERRISSPAETWITCEDRI 1246
DB 262 -----HQS----- 265

us-09-759-130b-381.rsp
-----
QY 1247 RVRGDTGTCGRVETWAGSMGTVCDDSDWDLAEAEVVCQQLGCGSALAALRDSFGOGTG 1306
DB 266 RLTCGADRCCEQVEVHFGRVWNTVCDSEWYPSKAVKVLCSLGCCTAVERPKGLPHSL-SG 324
QY 1307 TIWLDNRCKNGESFLWDCHAKPWQSDCKGHEKADGVRCS-GQSLKSL----- 1353
DB 325 RMYYS--CNGEELTSCNSWRFNNSNLCSQSLAARVLCSSASRLNLSLSTPEVPASVOTV 381
QY 1354 -----NASSGHLALILSSI-FGLLLL--VLFLFTWCRVKQKHLPLRVS 1396
DB 382 TIESVTVKIENKESRELMLLIPSLIVIGILLGLSLFIATIL-LRIKGYALPVWVN 437

RESULT 13
ID SPER_STRPU STANDARD; PRT; 532 AA.
AC P16264; 1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Egg peptide speract receptor precursor.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OC NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 477-489.
RX MEDLINE=89184581; PubMed=2538832;
RA Dangott L.J., Jordan J.E., Bellet R.A., Garbers D.L.;
RT "Cloning of the mRNA for the protein that crosslinks to the egg
peptide speract."
RL Proc. Natl. Acad. Sci. U.S.A. 86:2128-2132(1989).
CC -!- FUNCTION: RECEPTOR FOR THE EGG PEPTIDE SPERACT.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04518; AAA30078.1; -
DR PIR; A32751; A32751.
DR InterPro; IPR001190; Ssrcr_receptor.
DR Pfam; PF00530; SRCR_4.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00202; SR_4.
DR PROSITE; PS00420; SRCR_1; 4.
DR PROSITE; PS50287; SRCR_2; 4.
DR Receptor; Transmembrane; Glycoprotein; Signal; Repeat.
FT SIGNAL 1 30
FT CHAIN 31 532
FT DOMAIN 31 491
FT TRANSMEM 492 520
FT DOMAIN 521 532
FT DOMAIN 43 144
FT DOMAIN 153 257
FT DOMAIN 264 366
FT DOMAIN 382 485
FT CARBOHYD 78 78
FT CARBOHYD 115 115
FT CARBOHYD 459 459
FT SEQUENCE 532 AA; 57820 MW; 742533E095769CB8 CRC64;

Query Match
Best Local Similarity 29.1%; Score 520; DB 1; Length 532;
Matches 152; Conservative 66; Mismatches 208; Indels 96; Gaps 19;

```

QY 1 MMLPQNSWHIDFGRCCHQNLFSAVTCTILLNSCFLLISFNGFDLE-----LR 49  
Db 5 MMLQOYCV-----AACLVI---CIAISSVDDVGAQNYGREAVEGNIR 44  
QY 50 LVNGDGPSCGTVEVKFOGQWCTVCDGWNNTASTVVCVKQLGCPSPFAMFR---FGQAVTR 106  
Db 45 LIHGRTENEGSEVETIYHATIRGGVCDWWHMHENANVTCKLGFGPGARQFYRAYEGAHVT- 103  
QY 107 HGKILWLDVSCYGNESALWECQHEMG--SHNCHGEDGVNVCY--GEANLGLRLVDGNN 163  
Db 104 --TFWVVKMCLNETLEDYHPRYGRPWLCNAQWAGVECLPKDPEQSLRMLIGDVP 161  
QY 164 CSGRVEVKFOERWGTICDDGNLNTAAVVCVKQLGCPSPFSFSSGVNNSPA-----VLRPI 217  
Db 162 NEGTLFTWGDGAGSVCHTDFGPDGNVACRQMG-----YSRGVKSITKDGHFSTGPI 216  
QY 218 WLDDILCOGNELALWNCRHGRWGN-----HDC--SHNEDVTLCTCYDSDLELRVGGT 268  
Db 217 ILDAVDCGEGTEAHTEC-----NMPVTPYOHACPYTHNMWDVGVCKPNVEGDRLMDGS 270  
QY 269 NRCMGRVELKIQGRWGTVCCHKNNNAADVVCVKQLGCGTALHFAGLPHLOSGSDV----- 323  
Db 271 GPHEGRVEIWHDDAWGTICDDGWDADANVVCROAGYRGAVKASGF---KGEDFGFTWA 326  
QY 324 -VWLDGVSCGNSFLWDRHSRGTNFDCLHQNDVSVICSDGADLELRAD-----G 374  
Db 327 PIHTSFYMTGTGVEDRLICILRDGWTSHCYHVEDASVVCATDDDDTIEIEPKHTRVRIVG 386  
QY 375 SNNSCGRVEVRIHQWTTICDQNNKQALVVCVKQLGCPSPFSVFSRRAKPSNEA----- 428  
Db 387 MGQCGRVEVSLGNGWGRVCDPQDSHDEAKTVCYHAGYK---WGASRAAGSAEVSAPFDL 443  
QY 429 -RDWISISCTGNES-ALMDCTYDGRKAKTCTFRSRDAGVIC 468  
Db 444 EAPFIIDGTCVGENETLSQCKMKVSDMTCT-ATGDVGVC 484

## RESULT 14

MRCO\_MESAU  
ID MRCO\_MESAU STANDARD; PRT; 483 AA.  
AC QWUB9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Macrophage receptor MARCO (Macrophage receptor with collagenous structure).  
GN MARCO.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.  
OX NCBI\_TaxID=10036;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99242649; PubMed=10224280;  
RA Palecanda A., Paulauskis J., Al-Mutairi E., Imrich A., Qin G., Suzuki H., Kodama T., Tryggvason K., Koziel H., Kobzik L.;  
RT "Role of the scavenger receptor MARCO in alveolar macrophage binding of unopsonized environmental particles";  
RL J. Exp. Med. 189:1497-1506(1999).  
CC -1- FUNCTION: Bind Gram-positive and Gram-negative bacteria.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).  
CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SRCR DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF125191; AAD20360.1; -.  
DR InterPro; IPR000087; Collagen.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00530; SRCR; 1.  
DR Pfam; PF01391; Collagen; 4.  
DR PRINTS; PR00258; SPERACTRCPTR.  
DR PRODOM; PD000007; Collagen; 1.  
DR SMART; SM00202; SR; 1.  
DR PROSITE; PS00420; SRCR\_1; FALSE\_NEG.  
DR PROSITE; PS00287; SRCR\_2; 1.  
KW Collagen; Transmembrane; Receptor; Glycoprotein; Signal-anchor.  
FT DOMAIN 1 48 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 49 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
FT DOMAIN 70 483 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 148 383 COLLAGEN-LIKE.  
FT DOMAIN 389 483 SRCR.  
FT DISULFID 412 472 BY SIMILARITY.  
FT DISULFID 425 482 BY SIMILARITY.  
FT DISULFID 452 462 BY SIMILARITY.  
FT CARBOHYD 85 85 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 137 137 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 483 AA; 49621 MW; C38F18C46505FB1E CRC64;  
Query Match 3.7%; Score 298; DB 1; Length 483;  
Best Local Similarity 52.0%; Pred. No. 1.1e-13;  
Matches 53; Conservative 13; Mismatches 30; Indels 6; Gaps 2;

QY 1034 KRLRLVDGDSRCAGRVEIYHDFGWTICDDGWDLSDAHVVCQKLCGVAFAFNAVTSYAHGE 1093  
Db 387 KEVIRVGGTNR--GRAEIFYNNWGTICDDNNDNDAVFCRMLG-----YSSGKGFTFG 440  
QY 1094 GSGPIWLDLCTGESHLCWCPSPRGWGHDCRHKEDAGVIC 1135  
Db 441 GSGNIWLDVNCQGTEDSLWNCRNKNNWGSNCHNEDAGVEC 482

## RESULT 15

MTCO\_MOUSE  
ID MTCO\_MOUSE STANDARD; PRT; 518 AA.  
AC Q60754;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Macrophage receptor MARCO (Macrophage receptor with collagenous structure).  
GN MARCO.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95171455; PubMed=7867067;  
RA Elomaa O., Kangas M., Sahlberg C., Tuukkanen J., Sormunen R., Liakka A., Thesleff I., Kraal G., Tryggvason K.;  
RT "Cloning of a novel bacteria-binding receptor structurally related to scavenger receptors and expressed in a subset of macrophages";  
RL Cell 80:603-609(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99265975; PubMed=10331948;  
RA Kangas M., Brannstrom A., Elomaa O., Matsuda Y., Eddy R., Shows T.B., Tryggvason K.;  
RT "Structure and chromosomal localization of the human and murine genes for the macrophage MARCO receptor";  
RL Genomics 58:82-89(1999).  
CC -1- FUNCTION: Bind Gram-positive and Gram-negative bacteria.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).  
CC -1- TISSUE SPECIFICITY: Expressed in subpopulations of macrophages in the spleen and the medullary cord of lymph nodes.  
CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SRCR DOMAIN.

```
CC ----- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; UI8424; AAA68638.1; -
DR EMBL; AF128423; AADS1136.1; -
DR EMBL; AF127927; AADS1136.1; JOINED.
DR EMBL; AF127928; AADS1136.1; JOINED.
DR EMBL; AF128169; AADS1136.1; JOINED.
DR EMBL; AF128170; AADS1136.1; JOINED.
DR EMBL; AF128171; AADS1136.1; JOINED.
DR EMBL; AF127601; AADS1136.1; JOINED.
DR EMBL; AF127602; AADS1136.1; JOINED.
DR EMBL; AF128419; AADS1136.1; JOINED.
DR EMBL; AF128420; AADS1136.1; JOINED.
DR EMBL; AF128421; AADS1136.1; JOINED.
DR EMBL; AF128422; AADS1136.1; JOINED.
DR MGD; MGI:1309998; Marco.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001190; Sctr_receptor.
DR Pfam; PF00530; SRCR; 1.
DR Pfam; PF01391; Collagen; 5.
DR PRINTS; PR00258; SPERACTRCPTR.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS0287; SRCR_2; 1.
KW Collagen; Transmembrane; Receptor; Glycoprotein; Signal-anchor.
FT DOMAIN 1 48 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 49 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 70 518 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 149 418 COLLAGEN-LIKE.
FT DOMAIN 423 518 SRCR.
FT DISULFID 446 507 BY SIMILARITY.
FT DISULFID 459 517 BY SIMILARITY.
FT DISULFID 487 497 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 518 AA; 52730 MW; B09E7601ECA3637 CRC64;

Query Match 3.6%; Score 290.5; DB 1; Length 518;
Best Local Similarity 49.5%; Pred. No. 4.2e-13;
Matches 51; Conservative 16; Mismatches 31; Indels 5; Gaps 2;

QY 1034 KRLRLVDGSRGAGVEIYHDFGWTICDDGWDLSDAHVVCKLGGGVAFNATVSAHFGE 1093
DB 421 QVRIRMGGTNR--GRAEVYNNWGTICDDNDNDATVFCRMIGYS---RGRALSSYGG 475

QY 1094 GSGPIWLDLNCGTGTFESHLWCQPSRGWGQDCHRHKEDAGVICS 1136
DB 476 GSGNIWLDNVNCGTENSILWDCSKNSGHNVCVHNEDAGVECS 518
```

Search completed: May 12, 2003, 06:37:50  
Job time : 28 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:36:46 ; Search time 24 Seconds  
(without alignments)  
1781.313 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSHIDFGRCCHQ.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata1/iaa/5A\_COMB.pep.\*

2: /cgn2\_6/ptodata1/iaa/5B\_COMB.pep.\*

3: /cgn2\_6/ptodata1/iaa/6A\_COMB.pep.\*

4: /cgn2\_6/ptodata1/iaa/6B\_COMB.pep.\*

5: /cgn2\_6/ptodata1/iaa/PCTUS\_COMB.pep.\*

6: /cgn2\_6/ptodata1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2435	29.9	1785	4	US-09-341-587-3
2	1048.5	12.9	1290	1	US-08-470-350B-2
3	718.5	8.8	666	4	US-09-341-587-1
4	691.5	8.5	346	3	US-09-034-916-2
5	629	7.7	754	4	US-09-276-400-8
6	629	7.7	754	4	US-09-448-076-8
7	629	7.7	754	4	US-09-702-572-8
8	622.5	7.6	753	4	US-09-276-400-2
9	622.5	7.6	753	4	US-09-448-076-2
10	622.5	7.6	753	4	US-09-702-572-2
11	607.5	7.5	774	4	US-09-276-400-7
12	607.5	7.5	774	4	US-09-448-076-7
13	607.5	7.5	774	4	US-09-702-572-7
14	462	5.7	100	3	US-09-034-916-15
15	432	5.3	100	3	US-09-034-916-17
16	412	5.1	100	3	US-09-034-916-13
17	347	4.3	100	3	US-09-034-916-11
18	346	4.3	102	3	US-09-034-916-16
19	338	4.2	100	3	US-09-034-916-10
20	337.5	4.1	101	3	US-09-034-916-9
21	326.5	4.0	95	3	US-09-034-916-7
22	311.5	3.8	585	1	US-08-477-674-10
23	311.5	3.8	585	1	US-08-473-791-10
24	311.5	3.8	585	2	US-08-316-714-10
25	311.5	3.8	585	3	US-08-473-673-10
26	311	3.8	100	3	US-09-034-916-14
27	296	3.6	102	3	US-09-034-916-5

28	290.5	3.6	489	2	US-08-794-795-7	Sequence 7, Appli
29	290.5	3.6	489	4	US-09-249-200-7	Sequence 7, Appli
30	290.5	3.6	518	1	US-08-392-367B-2	Sequence 2, Appli
31	290.5	3.6	518	3	US-08-893-467A-2	Sequence 2, Appli
32	290	3.6	451	1	US-08-154-365-2	Sequence 2, Appli
33	288.5	3.5	495	2	US-08-794-795-2	Sequence 2, Appli
34	288.5	3.5	495	4	US-09-249-200-2	Sequence 2, Appli
35	288.5	3.5	520	2	US-08-794-795-6	Sequence 6, Appli
36	288.5	3.5	520	4	US-09-249-200-6	Sequence 6, Appli
37	288	3.5	451	1	US-08-453-117-2	Sequence 2, Appli
38	288	3.5	451	2	US-08-948-222-2	Sequence 2, Appli
39	288	3.5	451	2	US-08-973-145-2	Sequence 2, Appli
40	288	3.5	451	4	US-09-276-400-10	Sequence 10, Appli
41	288	3.5	451	4	US-09-448-076-10	Sequence 10, Appli
42	288	3.5	451	4	US-09-702-572-10	Sequence 10, Appli
43	288	3.5	451	5	PCT-US96-08081-2	Sequence 2, Appli
44	278	3.4	101	4	US-09-518-046-16	Sequence 16, Appli
45	273	3.4	101	4	US-09-518-046-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1

US-09-341-587-3

; Sequence 3, Application US/09341587

; Patent No. 6346606

; GENERAL INFORMATION:

; APPLICANT: Mollenhauer, Jan

; TITLE OF INVENTION: Protein Containing an SRCR Domain

; FILE REFERENCE: 4121-108

; CURRENT APPLICATION NUMBER: US/09/341.587

; CURRENT FILING DATE: 1999-08-31

; EARLIER APPLICATION NUMBER: PCT/DE98/000096

; EARLIER FILING DATE: 1998-01-09

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 3

; LENGTH: 1785

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-341-587-3

Query Match 29.9%; Score 2435; DB 4; Length 1785;  
Best Local Similarity 36.1%; Pred. No. 3.9e-206;  
Matches 533; Conservative 155; Mismatches 415; Indels 374; Gaps 28;

Qy	43	GTD--LELRVNGDPCSGTVEVKFQGMGTVCDDGNWTTASTVVCQKLGCPFSFAM---	97
Db	95	GSDSGLALRLVNGDRCQGRVEILYRGSGTVCDDSDTNDANVVCRLGCGWANSAPGN	154
Qy	98	FRFQAVTRHCKIWLDDVSCYGNESALWECOHREWSHNCYHGDEGVNCY-----	148
Db	155	AWFQGG---SGPIALDDVRCSGHESYLWSCP HNGHGWLTHNCGHGSDAGV	211
Qy	149	-----GEANLGLRLVDGNNSCGRVEVKFQGRVETCTDCDGNLNTAAV	192
Db	212	PESWPVRISSPPTGESSLSLRLVNGDRCRCRVEILYRGSGTVCDDYTDNDANV	271
Qy	193	CRLGCPFSSTISSGVNPAVLRLPILWDDILCQGNELALNCRHRGNHDCSHNEDVTL	252
Db	272	CRLGCGWANSAPGACGFGGSGPIVLLDDVRCSGHESYLWSCP HNGHGWLTHNCGH	331
Qy	253	TCY-----DSSDLRLVGTNRCMGRVLEKIQGRWCTVCHHK	290
Db	332	ICSAQSRPTSPDTPWTPTSHASTAGPESLSLRLVNGDRCRCRVEILYRGSGTVCDD	391
Qy	291	WNAADVCKQLCGCTALHFAGLPHLQSGDVVWLGVSCSGNESFLWDCRHSGTGNFD	350
Db	392	WDTSDANVVCRLGCGWATAPGNARFGQSGPIVLLDDVRCSGYESYLWSCP HNGH	451
Qy	351	CLHQNDYSVICSQDGLLELRADSGNNSCGRVEVRIHEQWMTICDQNKWNEQALVVCQ	410

Db 452 CQHSDEAGVCSAA-----HSWT-----470  
QY 411 GCPFSVFGSRRAPNEARDIWINSICTGNESALWDCTYDGKAKRTCFRRSDAGVICSD 470  
Db 471 -----PSPDLPTITLPASTVGSESS-----491  
QY 471 KADLDRLVGAHSPCYGRLEVYKQGEWGTCHDRMSTRNAAVVCQKLGCGKPMHVFGMTY 530  
Db 492 ---LALRVNGGDRCGRVEVLYQGSWGTCDSDWTDNDANVVCQPCGCGWAMSAPGNAR 548  
QY 531 FKEASPTLWDDVSCIGNESINWDCHEHSGKHNCHVREDVIVTCSGDA-----TW- 581  
Db 549 FQGGSPVLDDVRCSGHESYFWSCPHNGWLSHNCGHSDAGVICSASQSRPTPSPTWP 608  
QY 582 -----GURLVGGNRCGRLEVYFOGRWGTVCDDGWNSKAAAVVCSOLDPCS 628  
Db 609 TSHASTAGSESSALRLVNGGDRCGRVEVLYRGSWGTVCDDYWDNDANVVCRLGCGW 668  
QY 629 STIIGMLGNASTGYKILWDDVSCDGSDDLWSCNRSNGNNDCHSHSEDDVGIVCS-----683  
Db 669 AMSAPGNARFGGSPVLDDVRCSGHESYFWSCPHNGWLSHNCGHSDAGVICSASQSQ 728  
QY 684 -----DASOMELRVGSSRCACGVVNVGGVAILCANGWGMNIAEV 726  
Db 729 PTPSPDPTWPTSHASTAGSESSALRLVNGGDRCGRVEVLYRGSWGTVCDDYWDNDANV 788  
QY 727 VCRQLECGSAIRVREPHTERTLHILMSNGCTGEASLWDCIRWENKQATACHLNMESAS 786  
Db 789 VCRQLCGGWATAPGNARFGGSPVLDDVRCSGHESYFWSCPHNGWLSHNCGHSDAG 848  
QY 787 LICSAHR-OPRLVGADMPGCGRVEVYKHDATWRSVCDSDFSLHAANVLCRELNCDAISLS 845  
Db 849 VICASASQP-----TPSPDPTWPT-----867  
QY 846 VGDHFGKGNLWAEKFCQEGSETHALCPVQHPEDTCHSREYVVCVRYTDLVNG 905  
Db 868 -----SRASTAGSESTAL-----RLVNG 886  
QY 906 KSCQGOVEINVLHGWSLCTDHPDEARVLCRLSGTALSTGGKYIGRSYRVNMGH 965  
Db 887 GDRGRVLEVLYQSGWGTVCDDYWDNDANVVCRLGCGWAMAFGNAQFGGSGPIVLD 946  
QY 966 RFHCLGNESLLDNQNTVLGAPPCITGNTVSVICTGSLTQ--PLPCLANVSDPYLSAVP 1023  
Db 947 DYRCSGHESYLSWCPHNGWLSHNCGHSDAGVICSASQSPREDTTLTTLNLPALTVGS 1006  
QY 1024 EGSALICLEDKRLVLDGSDRCAGRVEYHDCGFWGTICDDGWLSDAHVVCOKLGGVAF 1083  
Db 1007 ESSL-----ALRVNGGDRCGRVEVLYRGSWGTVCDDSDWTDNDANVVCRLGCGWAM 1059  
QY 1084 NATVSAHFEGSGPIWLDLNCITGESHLMQCPSRGWGQHCRRKEDAGVICSEFTALRL 1143  
Db 1060 SAPGNARFGGSPVLDDVRCSGHESYFWSCPHNGWLSHNCGHSDAGVICS---ATQI 1116  
QY 1144 YSET-----ETESACGRLEVY-NGTWGS-----1166  
Db 1117 NETTTDWWHTTTTARPSNCGGL--FYASGTFSPPSYPAYPNNAKCVMEIEVNSGY 1174  
QY 1167 ---VGRNIT-----TTAIAIGV---ROLGCGENGWVSL---AP 1196  
Db 1175 RNLGFSNLKLAHNCSPDYVEIFDGSLSNLLGKICNDTRQIFTSYNYMTHTFRSD 1234  
QY 1197 LSKTGSFPM-WYDDIOCPKTHISIWQCLAPWERRISSPAETWITCEDRIRVRGDTFC 1255  
Db 1235 ISFQNTGFLAWYN-----SPSDATL-----RLVNLNSSYGLC 1267  
QY 1256 SGRTVTHAGSWGTVCDDSDWDLAEAEVVCQQLGCGSALAAALRDASEGQGTGTIWLDDMRC 1315  
Db 1268 AGRVEIYHGGTGTVCDDSWTIQAEAEVVCRLGCGRAVSALGNAYFGSGSPITLDDVBC 1327  
QY 1316 KGNESFLWCPCHAKPMQSCDGHKEDAGVRCSGSLKS 1352  
Db 1328 SGTESTLMQCRNGWFSHNCNHRDAGVICSNGHLST 1364

## RESULT 2

US-08-470-350B-2  
; Sequence 2, Application US/08470350B  
; Patent No. 5684126  
; GENERAL INFORMATION:  
; APPLICANT: Li, Xiao  
; APPLICANT: Snyder, Solomon H  
; TITLE OF INVENTION: Ebnerin: A Secreted von Ebner's Gland  
; TITLE OF INVENTION: Protein Associated with Taste Buds  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff, Ltd.  
; STREET: 1001 G Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/470,350B  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wolffe, Susan A  
; REGISTRATION NUMBER: 33,568  
; REFERENCE/DOCKET NUMBER: 01107.48790  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1290 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-470-350B-2

Query Match 12.9%; Score 1048.5; DB 1; Length 1290;  
Best Local Similarity 29.7%; Pred. No. 1.le-83;  
Matches 289; Conservative 114; Mismatches 290; Indels 281; Gaps 34;

QY 286 VCHHKNNAAADVCKQLGCGTALHFAGLPHLQSGSDVVWLDGSCSGNESFLWPCRHS 345  
Db 1 MCDSDWINDANVVCRLGCGWALSAPGSAQFGQLGPIVLDVACRGHEAYLWCSHRG 50  
QY 346 TVNFDCLHNDVSVICSDGADLELRLADGNNCSGRVEVRIHQWTTICDQNKNEQALV 405  
Db 61 WLSHNCGHQEDAGVICSD-----SQTSSPTG-----WV---NPGGTNDVIY 100  
QY 406 VCKQLGCPFSVFGSRRAPKPSNEARDIWINISCTGNESALWDCTYDGKAKRTCFRRSDAG 465  
Db 101 DTQET-----TETSQTSSPTDWM-----NHGTTINDVIYD---TOETTEGTDSD 142  
QY 466 VICSADKADLDRLVCAHSPCYGRLEVYKQGEWGTCHDRMSTRNAAVVCQKLGCGKPMH 525  
Db 143 -----LAVRLVNGGDRCGRVEVLYRGSWGTVCDDSDWINDANVVCRLGCGWALS 194  
QY 526 FGMTYFKEASGPIWLDVSCIGNESINWDCHEHSGKHNCHVREDVIVTCS-----576  
Db 195 PCSAQFGGSGSIVLDDVACRGHEAYLWCSHRGLSHNCGHQEDAGVICSYSQTSSPTP 254  
QY 577 -----GDATWG-----LRLVGGNRCGRLEVYFGGRWG 605  
Db 255 DSQTSSTPTGWNPNPGTNDVSYGPEQTDTATDSGLAVRLVNGGDRCGRVEVLYQGSWG 314  
QY 606 TVCDDGWNSKAAAVVCSOLDPCSIIIGMLGNASTGYKILWDDVSCDGSDDLNMESAS 665

Db 315 TVCDDSDMTKDANVVCRLVCGWALSAPGSAHFQGGSGSIVLDDVACTHEAYLWSCSHR 374  
Qy 666 GWGNNDGSHSDVGVICSDAS-----DM-----ELRLVGGSSRCAGKVEVNVQGA 712  
Db 375 GWSHNCCHHEDAGVICSDAQTSFTWPDWPTTPTTDDWTKYSSVPTQFPTI- 433  
Qy 713 ILCANGMGNTAEVVCRLQECGSAIRVS-----REPHTFTERTLHILMSNGCTGGEASLWD 768  
Db 434 ---ADMWTPSPYEV-----CGLLTLTPYQGFSSPYPGS-----YPNNARC-----LW- 474  
Qy 769 CIRWENKOTACHLNMESALICSARQPLVCAADPCSGRVEVKHADTWSRSCDSESL-- 826  
Db 475 -----KIFVSSMNRVTWFTDQLEGG-----CNDYDILVF 505  
Qy 827 ----HAANVLCRELNCGDAISLSVDHFGK-----NGLTWAKEFOCE--- 865  
Db 506 DGPENNSLIAR-----VCDGFNGSFTSTQNFMSVVFITDGSVTRRGFOADYS 554  
Qy 866 ----GSETHALCPTV-----QHPEPTC-----IHSREVGVCSRYTDVRLVNGKSOCDGQV 913  
Db 555 TPISTSTSTPTTFTPIVTDWMTTPSPVTCGGLLTLPYQGFSSPY----- 598  
Qy 914 EINVLGHWSLCDTHWDPEDARVLCR-----QLSCGT-----ALSTTGCK 953  
Db 599 -----YPGSY-----PNNARCWKIFVPSMNRVTWFTDQLEGGCNDYILGFDGPE 646  
Qy 954 YIGERSVRVMGHRFCHLGNESLLDNCOMTVLGAAPCIHNTVSV--ICTGSLTOPLF--- 1008  
Db 647 YNLSLIAR-----CDGNSGFTSTQ-----NFMVVFITDGSVTRRGFOAD 688  
Qy 1009 ----PCLANVSDPYLSAPEG--SALICLEDKRLRLVDGSRACGRVEIYHDGFWGTICD.1062  
Db 689 YYSTPIRTSTPTTFTPIITGNDSSLV-----LRLVNGTNRCEGRVEILYRGSWVPACD 742  
Qy 1063 DCWLDSDAHVVCQKLGCGVAFNATVSAHFEGGSGPIWDLNLCNCTGESHLMQCPSPRGWGQ 1122  
Db 743 DSWINDANVVCRLQCGSALSAPGNAWFGGSLVLLVDVYSCGYESHLNCRHPGLV 802  
Qy 1123 HCRHKEDAGVICS 1136  
Db 803 HNCRVEDAGVICS 816

RESULT 3  
US-09-341-587-1  
; Sequence 1, Application US/09341587  
; Patent No. 6346606  
; GENERAL INFORMATION:  
; APPLICANT: Mollenhauer, Jan  
; TITLE OF INVENTION: Protein Containing an SRCR Domain  
; FILE REFERENCE: 4121-108  
; CURRENT APPLICATION NUMBER: US/09/341,587  
; CURRENT FILING DATE: 1999-08-31  
; EARLIER APPLICATION NUMBER: PCT/DE98/00096  
; EARLIER FILING DATE: 1998-01-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 666  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-341-587-1  
  
Query Match 8.88; Score 718.5; DB 4; Length 666;  
Best Local Similarity 35.88; Pred. No. 6.7e-55;  
Matches 162; Conservative 63; Mismatches 118; Indels 109; Gaps 13;  
  
Qy 971 GNEISLDNCOMTVLGAAPCIHNTVSVICTGSLTQ--PLFFCLANVSDPYLSAVPEGSAL 1028  
Db 1 GHESYLSNCPHNGWLSHNCGHEDAGVICSAAQSQSPRPDTLTLNLPALTVGSESL- 59  
Qy 1029 ICLEDKRLRLVDGSRACGRVEIYHDGFWGTICDDGWDLSDAHVVVCQKLGCGVAFNATVS 1088  
|||||:| ||| ||||:| ||| |||| || ||| |||| || ||| ||||:|||||:| |

Db 60 -----ALRLVNGDRGRGRVEVLYRGSWGTVCDDSWDNDANVVCRLQCGWAMSAPGN 113  
Qy 1089 AHFEGGSGPIWDLNLCNCTGESHLMQCPSPRGWGOHDCRHKEDAGVICSEFTALRYSET- 1147  
Db 114 ARFGGSGPIWDDVRCSGNESYLSWSCPCHKGLWLNCHGHHEDAGVICS---ATQINSTTT 170  
Qy 1148 -----ETESCAGRLVEFY-NGTWGS-----VG 1168  
Db 171 DMWHPTTTTTPARPSSNCGGFL--FYASGTFSSPSYPAYIPNNAKVMEIEVNSGYRINLG 228  
Qy 1169 RRNI-----TTAIGIVC-----RQLCGGNGVVSLL---APLSKTG 1201  
Db 229 PSNLEAHNHCSDFYVEIFDGLNSLLKGLKICNDTRQIFTSSYNRMTIHFSDISFQN 288  
Qy 1202 SGFM-WVDDIOCPKTHISIMQCLSAPWERRISSPAEETWITCEDRIRVRGGDTSCGRVE 1260  
Db 289 TGFLAWYN-----SFFSDATL-----RLVNLNSSYGLCAGRVE 321  
Qy 1261 IWHAGSMGTVCDDSWDLAEAEVVCQQLCGSALALDRDASFQGGTGTIWLDDMRCKGNES 1320  
Db 322 IYHGTGTWGTVCDDSWTIOEAEVVCRLQCGRAVSALGNAYFGSGGPIITLDDVECSGTES 381  
Qy 1321 FLWDCHAKPWQSCDGHKEDAGVRCSGOSLAKS 1352  
Db 382 TLWQCRNRGWFSHNCNHRDAGVICSGNHLST 413

RESULT 4  
US-09-034-916-2  
; Sequence 2, Application US/09034916  
; Patent No. 6046314  
; GENERAL INFORMATION:  
; APPLICANT: GEBE, JOHN A.  
; APPLICANT: SIADAK, ANTHONY W.  
; APPLICANT: ARUFFO, ALEJANDRO A.  
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR  
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBINS & ASSOCIATES  
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
; CITY: MENLO PARK  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,916  
; FILING DATE: 04-MAR-1998  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/039,956  
; FILING DATE: 06-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5998-0020  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 325-7812  
; TELEFAX: (650) 325-7823  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 346 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-034-916-2

```

Query Match      8.5%; Score 691.5; DB 3; Length 346;
Best Local Similarity 41.3%; Pred. No. 6.1e-53;
Matches 148; Conservative 40; Mismatches 139; Indels 31; Gaps 9;

QY      21  LFSAVVTCILLNSC-----FLISSEFNGTDLRLVLVNGDGPCSGVVEYKFGOWGTVCDDG 76
Db      3  LFS-----LILAICTRPGFLASPSG-----VRLVGLGHRCEGRVVEYKQGWGTVCDDG 51

QY      77  WNTTASTVVCVKQLCGPFSFAM--FREGQAVTRHGKIWLDDVSYGNSALSECQHREWG 133
Db      1  WNTTASTVVCVKQLCGPFSFAM--FREGQAVTRHGKIWLDDVSYGNSALSECQHREWG 133

QY      134 SHNCYHGDEDVGVNC-----YGEANLGLRLVDGNNSSCGRVEYKFORRWGTICDDGNLNL 187
Db      1  SHNCYHGDEDVGVNC-----YGEANLGLRLVDGNNSSCGRVEYKFORRWGTICDDGNLNL 187

QY      188 TAAVVCROLGCPSPFISSGVYVNSPAVLRLPIWLDDILCOGNELALWNCRRHGWGNDHCSH 246
Db      1  TAAVVCROLGCPSPFISSGVYVNSPAVLRLPIWLDDILCOGNELALWNCRRHGWGNDHCSH 246

QY      170 AAKVVCROLGCGRAVLTKQCNKHAYGRKPWLWSQMSGREATLQDCPSGFWGNKNTCNH 229
Db      1  AAKVVCROLGCGRAVLTKQCNKHAYGRKPWLWSQMSGREATLQDCPSGFWGNKNTCNH 229

QY      247 NEDVTLTCYDSSDLELRLVGVTNRCMGRVELKIOGRWGTVCHHKKNWNAADAVVCVKQLGCG 306
Db      1  NEDVTLTCYDSSDLELRLVGVTNRCMGRVELKIOGRWGTVCHHKKNWNAADAVVCVKQLGCG 306

QY      230 DEDTWECEDPFD--LRLVGDNLCSGLEVYLHKGVWGSVCDNNGEKEDQVVCVKQLGCG 287
Db      1  DEDTWECEDPFD--LRLVGDNLCSGLEVYLHKGVWGSVCDNNGEKEDQVVCVKQLGCG 287

QY      307 TAL--HFAGLPHLQSGSDVYVWLGDVCSGNSBSFLWDCRHSGTSTVNFDCQLHQNDSVICS 362
Db      1  TAL--HFAGLPHLQSGSDVYVWLGDVCSGNSBSFLWDCRHSGTSTVNFDCQLHQNDSVICS 362

QY      288 KSLSPSFRDRKCYGPGVGRVWLQVNRVCSGSEBOSLEOCOHREHFGWGHDCPHOEDVAVICS 345
Db      1  KSLSPSFRDRKCYGPGVGRVWLQVNRVCSGSEBOSLEOCOHREHFGWGHDCPHOEDVAVICS 345

```

## RESULT 5

```

US-09-276-400-8
; Sequence 8, Application US/09276400
; Patent No. 6140056
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276,400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 754
; TYPE: PRT
; ORGANISM: murine lysyl oxidase-related protein
; US-09-276-400-8

```

[illegible]

```

Db      340 QASVVCPELGFGTAREALSCARMQCGMALHLSVEVRCSGEPELSWRCPSKNITAECDCH 399
      11 111 :11 111 : : : : 111 : 111 : 111 : 111 : 111 :
QY      354 ONDVSVTCS---DGADLELERLADGNNCSGRVEVRI---HEQWWTICDQWKNQALVY 406
      11 : : :111 : : 111111 : : : : 11 : 11 : 11 :
Db      400 SDAGVRCNLPYTGVTETKRIILSGGRSRYEGREYVOIGIPGHLRWGLICGDDWGTEAMVA 459
      11 : : :111 : : 111111 : : : : 11 : 11 : 11 :
QY      407 KQLGCFSEFVSGSRRRA--KPSNEARDIWINISICTGNEALSALWDCDTCYDGKAKRTCFR---R 461
      111111 : : : : : : : : : : : : : : : : : : : : : :
Db      460 CRQLGLGYANHGLQETWYWDGNGVTEVMVSGVRCGSELSLNLQCAHH--SSHITCKKTGTR 518
      111111 : : : : : : : : : : : : : : : : : : : : : :
QY      462 SDAGVICSDKADLDLRLVGA 481
      111111 : 111 :
Db      519 FTAGVICSETAS-DLLLSHA 537
      111111 : 111 :

RESULT 6
US-09-448-076-8
; Sequence 8, Application US/09448076
; Patent No. 6300092
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran et al.
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN
; FILE REFERENCE: MNI-073CP
; CURRENT APPLICATION NUMBER: US/09/448,076
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/117,580
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 09/276,400
; EARLIER FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 754
; TYPE: PRT
; ORGANISM: murine lysyl oxidase-related protein
US-09-448-076-8

```

[illegible]

QY	462	SDAGVICSADKADLDRLVGA	481
DB	519	FTAGVICSETAS-DLLLSA	537
 RESULT 7 US-09-702-572-8 ; Sequence 8, Application US/09702572 ; Patent No. 6391602 ; GENERAL INFORMATION: ; APPLICANT: Khodadoust, Mehran ; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND ; FILE REFERENCE: MNI-073 ; CURRENT APPLICATION NUMBER: US/09/702,572 ; PRIOR FILING DATE: 2000-10-31 ; PRIOR APPLICATION NUMBER: 09/276,400 ; PRIOR FILING DATE: 1999-03-25 ; NUMBER OF SEQ ID NOS: 10 ; SOFTWARE: PatentIn Ver. 2.0 ; SEQ ID NO 8 ; LENGTH: 754 ; TYPE: PRT ; ORGANISM: murine lysyl oxidase-related protein ; US-09-702-572-8			
 Query Match            7.7%; Score 629; DB 4; Length 754; Best Local Similarity 31.0%; Pred. No. 6,7e-47; Matches 155; Conservative 72; Mismatches 183; Indels 90; Gaps 16;			
QY	56	PCSGTVEVKFGQNGTCDDGNTASTAVVCVKQLGCPFSFA-----MFRFGQAVTRHGKI	110
DB	54	PYEGRVEIQRAGEWGTICDDDTLQAHHVLCRELG--FTEATGWTHSAKYGPCT---GRI	108
QY	111	WLDDVSCYGNESALWECQHREWSHNCYHGEDVGNYC-----YGEANL-----	153
DB	109	WLDNLSCRGTEGSVTECASRGNSDCTHDEDAGVICKDQRLPGFSDSNVIEVEHQLOVE	168
QY	154	-----GLRLVDGNNSCGRVVEKFQERWGTCDDGWNLTAAVVCROLCGPSS--	201
DB	169	EVRLRPAVENGRPLP---VTEGLVEVRLPEGWSQVCDKGNSAHNSHVVCMLGFPGPKR	225
QY	202	-----FISSGVVNSPAVLPIWLDLILCOGNELALNCRHRGNGHD---CSHNEDVT	251
DB	226	VNMAFYRLAQKKQHS-----FGLSHVACVCTEAHLSLCSLEFYRANDTTTRCSGPNPV	279
QY	252	LTCY-----DSDLRLVGGTNRCMRGRVELKIQRWGTVCCHHKWN	293
DB	280	VSCVGLPLYATFTGKKOOHSPKEARVRLKGGAHQEGREVLEVLAGTWGTVCDRKWDL	339
QY	294	AAADVCKOLCGTALHFAGLPQLSQSDVYLDCVSCGSEFLWDCRHSGTVNFDCILH	353
DB	340	QAASVCPCLGFTAREALSARGMQGGAITHLSEVRCSQGPEPSLWRCPSKNITAEBCSH	399
QY	354	QNDSVICS-----DGADLELRADGNNSCGRVEVRI----HEQWMTICDONKNNEQALVV	406
DB	400	SODAGVRNLPYTGVETKIRLSGGRSRYEGRVEQIGIPGHLRWGLICGDDWGTLEAMYA	459
QY	407	CKQLGCFPSVGSRA--KPSNEARDIWNISCTGNSALWDCITYDGAKKTCTPR---R	461
DB	460	CROGLGYANHGLQETWTYWDNGNTEVMVSVGRCVTSSELNQCAHH--SSHITCKTKTGTR	518
QY	462	SDAGVICSADKADLDRLVGA	481
DB	519	FTAGVICSETAS-DLLLSA	537
 RESULT 8 US-09-726-400-2 ; Sequence 2, Application US/09276400 ; Patent No. 6140056 ; GENERAL INFORMATION:			

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-448-076-2

Query Match  
Best Local Similarity 7.6%; Score 622.5; DB 4; Length 753;  
Matches 167; Conservative 71; Mismatches 202; Indels 95; Gaps 19;

QY 30 LLLNSCLFSSFNPT-----DLELRVN-GDGPCSGTVEVKFGQGWGTVCD 74  
DB 14 LLL-CLLSSCLGSPSTGPEKKAGSGLRFRLAGFPKPYEGRVEIQRAGEWGTICD 71  
QY 75 DGNWTTASTVVKQKGCFFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECOH 129  
DB 72 DDFTLQAAHILCRELG--FTEATGTHSAKYGPST--GRIWLDNLSCSGTEQSVTECAS 126  
QY 130 REMGSHNCHGEDVGVNC-----YGEANL-----CLRLVDGNN 162  
DB 127 RGWNSDCTHDEDAVICDQRLPGFSDSNVIEVEHHIQVEEVRIRPAVGWRPLP--- 183  
QY 163 SCGRVEVKFQERWGTICDDGNLNTAAVVCRLGCPSPSFSSGVVNSPAVL---PIWL 219  
DB 184 VTEGLVEVRLPDGWSQVCDKWSAHNSHVCGMLGFPSEKRVNAAFYRLAQROHSGFL 243  
QY 220 DDILCOGNELALWNRHGRGNHD---CSHNEDVTLTC-----YDSS----- 258  
DB 244 HGVACVGTGEAHLSCSLFERYANDTARCPGGPVPVYAASSGQKQKQKQKPG 303  
QY 259 DLELRVCGTNRGMRVELKIOGRWGTVCHHKWNNAAADVCKOLGCGTALHFAGLPHLQ 318  
DB 304 EARVRLKGAHPGEGRVELKASTGTWCDRKWDLHAASVVCRELFGSAREALSGARMG 363  
QY 319 SGSDVVWLDGVSCGNEFSLWDCRHSHTVNFDCLEHNDVSVICS---DGADLELRADGS 375  
DB 364 QMGAIHLSEVRCSEGLSLKCPKHNITAEDCSHQDAGVRCNLPYTGAEIRLSGR 423  
QY 376 NNCGRVEVRIHE-----QWWTICDNWKNQALVVKOLGCPFPFVSGSRA--KPSNEAR 429  
DB 424 SQHEGRVEVQIGGPGPLRWGLTCGDDWGTLEAMVACROGLGLGYANHGLQETWYDWSGNT 483  
QY 430 DIWINSISCTGNESALWDCYDGKAKRTCFR---RSDAGVICSDDKADLDLRLVGA 481  
DB 484 EYVMSGVRCGTGELSDQCAHHG--THITCKRTGTRTAGVICSETAS-DLLLHSA 536

RESULT 10  
US-09-702-572-2  
; Sequence 2, Application US/09702572  
; Patent No. 6391602  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran  
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND  
; FILE REFERENCE: MNI-073  
; CURRENT APPLICATION NUMBER: US/09/702,572  
; PRIOR FILING DATE: 2000-10-31  
; PRIOR FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 753  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-702-572-2

Query Match  
Best Local Similarity 7.6%; Score 622.5; DB 4; Length 753;  
Matches 167; Conservative 71; Mismatches 202; Indels 95; Gaps 19;

QY 30 LLLNSCLFSSFNPT-----DLELRVN-GDGPCSGTVEVKFGQGWGTVCD 74  
DB 14 LLL-CLLSSCLGSPSTGPEKKAGSGLRFRLAGFPKPYEGRVEIQRAGEWGTICD 71

QY 75 DGNWTTASTVVKQKGCFFSFA-----MFRFGQAVTRHGKIWLDDVSCYGNESALWECOH 129  
DB 72 DDFTLQAAHILCRELG--FTEATGTHSAKYGPST--GRIWLDNLSCSGTEQSVTECAS 126  
QY 130 REMGSHNCHGEDVGVNC-----YGEANL-----CLRLVDGNN 162  
DB 127 RGWNSDCTHDEDAVICDQRLPGFSDSNVIEVEHHIQVEEVRIRPAVGWRPLP--- 183  
QY 163 SCGRVEVKFQERWGTICDDGNLNTAAVVCRLGCPSPSFSSGVVNSPAVL---PIWL 219  
DB 184 VTEGLVEVRLPDGWSQVCDKWSAHNSHVCGMLGFPSEKRVNAAFYRLAQROHSGFL 243  
QY 220 DDILCOGNELALWNRHGRGNHD---CSHNEDVTLTC-----YDSS----- 258  
DB 244 HGVACVGTGEAHLSCSLFERYANDTARCPGGPVPVYAASSGQKQKQKQKPG 303  
QY 259 DLELRVCGTNRGMRVELKIOGRWGTVCHHKWNNAAADVCKOLGCGTALHFAGLPHLQ 318  
DB 304 EARVRLKGAHPGEGRVELKASTGTWCDRKWDLHAASVVCRELFGSAREALSGARMG 363  
QY 319 SGSDVVWLDGVSCGNEFSLWDCRHSHTVNFDCLEHNDVSVICS---DGADLELRADGS 375  
DB 364 QMGAIHLSEVRCSEGLSLKCPKHNITAEDCSHQDAGVRCNLPYTGAEIRLSGR 423  
QY 376 NNCGRVEVRIHE-----QWWTICDNWKNQALVVKOLGCPFPFVSGSRA--KPSNEAR 429  
DB 424 SQHEGRVEVQIGGPGPLRWGLTCGDDWGTLEAMVACROGLGLGYANHGLQETWYDWSGNT 483  
QY 430 DIWINSISCTGNESALWDCYDGKAKRTCFR---RSDAGVICSDDKADLDLRLVGA 481  
DB 484 EYVMSGVRCGTGELSDQCAHHG--THITCKRTGTRTAGVICSETAS-DLLLHSA 536

RESULT 11  
US-09-276-400-7  
; Sequence 7, Application US/09276400  
; Patent No. 6140056  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran  
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND  
; FILE REFERENCE: MNI-073  
; CURRENT APPLICATION NUMBER: US/09/276,400  
; CURRENT FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 774  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-276-400-7

Query Match  
Best Local Similarity 7.5%; Score 607.5; DB 4; Length 774;  
Matches 196; Conservative 86; Mismatches 258; Indels 163; Gaps 28;

QY 260 LELRLVGTNR-CMGRVELKIQGRWGTVCHHKWNNAAADVCKOLGCGTALHFAGLPHLQ 318  
DB 56 IOLRLAGOKRKHSEGRVEVYDQGWGTVCDDDFSIHAAHVVCRELGYVEAKSWTASSYG 115  
QY 319 SGSDVVWLDGVSCGNEFSLWDCRHSHTVNFDCLEHNDVSVICS---DGADLELRADG 374  
DB 116 KEGEPIWLDNLHCTGNEATLAACTNSGNGWGTCDKTEEDGVVVCSDKRIPIGFKEDNSLNG 175  
QY 375 SNNS-----GRVEVRTHEOWWTICDNWKNQALVVKQKQ 411  
DB 176 IENLNQVEDIRAILSTYKRTVPMEGYVEVKEGKTWKQICDKHWTAKNSRVVCGMFG 235  
QY 412 CP-----FSVFGSRRAPKPSNEARDIWNISICTGNESALWDC-----TYDGAKR 456  
DB 236 PFGERTYNTKVYKMFASRR-----KQYWPFSMDCTGTGTEAHISSCKLGPQVSLDPMKNV 289

236	Db	FPGBRTNTVYKMFASRR-----KORYWPPFSMDCTGTEAHLSCKLGPQVSLDPMKNV 289
457	Qy	TCFRRS DAGVTC-----SDKADLDL--RLVGAHSPCYGRLEVKYIQGEWT 499
290	Db	TCENGLPAVYSCVPGVQFSPDGPSRFRKAYKPEOPLVRLGGAYIGEGRVEVLKNGEWGT 349
500	Qy	YCHDRWSTRNAAVVCKOLGCGCKPMHVFGMTYFPKEASGPITWLDVYSCIGNESNIWDCEHSG 559
350	Db	VCDDKDWLVASVYCRRELFGSAAKEAVTGSRLQOGGIGPIHLNEIOQTGNEKSIIDCKFNA 409
560	Qy	WGKHNCVHREDVIVTCSGDATW---GLRLVGGSNRCSGRLEVYFQGR---WGTVCDDGW 612
410	Db	-ESQGCNHEEDAGYRCNTPAMGLQKQLRLNGRNPYEGRVEVLVERNGSLVWGMVCGQNW 468
613	Qy	NSKAAAVVCSQLDCPSSIIHGMLGNASTGYGKTW-----LDDVSCDGDSDSLW 660
469	Db	GIVEAMVYCRQL-----GLGFASNAFQETWYHGDVNSKNVSMYGVKCSGTLSLSA 519
661	Qy	SCRNSGWNDCSHSEDV-----GVICSD-ASDMEL--RLVGSSSRCAKGRVEVN 706
520	Db	HCRHDG-----EDVACPGGVQYAGVACSETAPDLVLNAEMVQQTITYLEDRPMFM 570
707	Qy	VQGVAGILCANGWGMNTAEVVYCRQLBCGSAIRVSRPHEPHTERLHILMS---NSGCT--- 760
571	Db	IQ-----CA-----MEENCLSAQAQTDPTTGYRLLRFRSSQIHNNGOSDFR 612
761	Qy	-----GGEASLW--DCIR--NEWKOTACH---LWNEASLICSAAHQPRLVGADMPCSGRVEVKH 812
613	Db	PKNGRPHAWIHDDCHRHYSHEVFTHYDLLNLNGTKVAEGHKASFCL-EDTECEGEDTQKNY 671
813	Qy	ADTWRSYVC-----DSDFSLHAANVLCRELNGC---DAISLSVGDH 849
672	Db	E-----CANFGDGIITMGCDWMYRHHIDICQWVDITDVPBGDY 708

## 219 00

```

RESULT 13
US-09-702-572-7
; Sequence 7, Application us/09702572
; Patent No. 6391602
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AN
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/702,572
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 7
; LENGTH: 774
; TYPE: FRT
; ORGANISM: Homo sapiens
US-09-702-572-7

```

[illegible]

Db 236 PFERTYNTKVKFASRR-----KQYWPFSMDCTEASHSSCKLGPQVSLDPMKNV 289  
QY 457 TCFRSDAGVIC-----SDKADLDL-RLVGAHSPCYGRLEVKYQGEWGT 499  
Db 290 TCENGLPAVSCVPCQVSPDGPSPFRKAYKPEQPLVRLGGAYIGEGRVEVLKNGEWT 349  
QY 500 VCHDRWSTRNAAVCKQLCGCKPMHVFMTYFKEASGPIWLDLDDVSCIGNESNIWDCEHSG 559  
Db 350 VCDKDWLVSASVCRLELFGSAAKAVTCSRLGQIGIPHLNEIOCTGNEKSIIDCKENA 409  
QY 560 WCKHNCVHREDVIVTCSDATW---GLRVGSGNRCGRLEVIYFQGR---WGTVCDDGW 612  
Db 410 -ESQCNHEEDAGVRCNTPAMGLQKRLNGRNPYEGRVEVLVNRGSLVMGMVCGQNW 468  
QY 613 NSKAAVCSQLDCPSSIIIGMGLGNASTGYGKIW-----LDDVSCDGDSEDLW 660  
Db 469 GIVEAMVVCRL-----GLGFASNAFOETWYHGDVNSKNVVMGVKCSGTFELSLA 519  
QY 661 SCRNWGNNDGSHSEV-----GVICSD-ASDMEL--RLVGGSSRCAGKVEVN 706  
Db 520 HCRHDG-----EDVACPOGGVOYGAGVACSETAPDLVLAEMVQOTIYLEDPRPFM 570  
QY 707 VQGAIVCILCANGWMNIAEWCRCLECGSAIRVSRPHFTERTILHMS-----NSGCT--- 760  
Db 571 LQ-----CA-----MEENCLASAAQTDPTTGYRLLRFSSQIHNNQSDFR 612  
QY 761 ---GGEASLW-DCIR-WEWKQTACH---LNMEASLICSNAHROPRLVGMADPCSGRVEVKH 812  
Db 613 PKNGRHWIWHDCRRHYSMEVFTYDILLNLNGTKVAEGHKASFCLEDECEGDIQXNY 671  
QY 813 ADTWSVC-----DSQFSLHAANVLCRELNGC--DAISLSVGDH 849  
Db 672 E-----CANFGDQGITNCWDMYRHDDICQWVDITDVPBGDY 708

RESULT 14  
US-09-034-916-15  
; Sequence 15, Application US/09034916  
; Patent No. 6046314  
; GENERAL INFORMATION:  
; APPLICANT: GEBE, JOHN A.  
; APPLICANT: SIADAK, ANTHONY W.  
; APPLICANT: ARUFFO, ALEANDRO A.  
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR  
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL  
; TITLE OF INVENTION: ANTIBODIES THERETO  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBINS & ASSOCIATES  
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
; CITY: MENLO PARK  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,916  
; FILING DATE: 04-MAR-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/039,956  
; FILING DATE: 06-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5998-0020  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 325-7812  
; TELEFAX: (650) 325-7823  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-034-916-17

TELEFAX: (650) 325-7823  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-034-916-15  
Query Match 5.7%; Score 462; DB 3; Length 100;  
Best Local Similarity 77.0%; Pred. No. 1.9e-33;  
Matches 77; Conservative 9; Mismatches 14; Indels 0; Gaps 0;  
QY 1037 RLVDGDSRCAGRVEIYHDGFWGTCTCDDGWLSDAHVVCQKLCGCVAFNATYSAHFGECSG 1096  
Db 1 RLVNGGRCAGRVEIYHSGSWGTCDDSDWLSDAHVVCRLGCGEAINATGSAHFGECSG 60  
QY 1097 PIWLDLCTCTESHWCQPSRGWGQCHDKRHEDAGVICS 1136  
Db 61 PIWLDKMGKESRIMQCHSHGWQQNCRHKEDAGVICS 100

RESULT 15  
US-09-034-916-17  
; Sequence 17, Application US/09034916  
; Patent No. 6046314  
; GENERAL INFORMATION:  
; APPLICANT: GEBE, JOHN A.  
; APPLICANT: SIADAK, ANTHONY W.  
; APPLICANT: ARUFFO, ALEANDRO A.  
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR  
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL  
; TITLE OF INVENTION: ANTIBODIES THERETO  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBINS & ASSOCIATES  
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
; CITY: MENLO PARK  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,916  
; FILING DATE: 04-MAR-1998  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/039,956  
; FILING DATE: 06-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5998-0020  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 325-7812  
; TELEFAX: (650) 325-7823  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-034-916-17

Query Match 5.3%; Score 432; DB 3; Length 100;  
Best Local Similarity 71.0%; Pred. No. 8.7e-31;  
Matches 71; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

Search completed: May 12, 2003, 06:44:15  
Job time : 31 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 12, 2003, 06:37:26 ; Search time 151 seconds  
(without alignments)  
885.519 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDFGRCCHQ.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92023710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	8138	100.0	1453	9	US-10-042-431-11
2	8138	100.0	1453	9	US-09-759-130B-381
3	7913	97.2	1413	9	US-10-042-431-13
4	7913	97.2	1413	9	US-09-759-130B-383
5	7426	91.3	1319	9	US-10-042-431-14
6	7426	91.3	1319	9	US-09-759-130B-384
7	3441.5	42.3	1151	9	US-09-977-577-13
8	3429	42.1	1116	9	US-09-977-577-10
9	3429	42.1	1156	9	US-09-977-577-12
10	3396.5	41.7	1149	9	US-09-977-577-11
11	3082	37.9	1436	9	US-10-042-431-78
12	3082	37.9	1436	9	US-09-759-130B-448
13	822.5	10.1	822	9	US-09-147-947-6
14	719.5	8.8	761	9	US-09-147-947-4
15	691.5	8.5	347	9	US-09-905-291A-148
16	691.5	8.5	347	9	US-09-902-853-148
17	691.5	8.5	347	9	US-09-907-824-148
18	691.5	8.5	347	9	US-09-907-841-148
19	691.5	8.5	347	9	US-09-904-011-148

20	691.5	8.5	347	9	US-09-906-742-148	Sequence 148, App
21	691.5	8.5	347	9	US-09-906-838-148	Sequence 148, App
22	691.5	8.5	347	9	US-09-907-613-148	Sequence 148, App
23	691.5	8.5	347	9	US-09-907-942-148	Sequence 148, App
24	691.5	8.5	347	9	US-10-227-884-114	Sequence 114, App
25	691.5	8.5	347	9	US-09-904-820-148	Sequence 148, App
26	691.5	8.5	347	9	US-09-904-859-148	Sequence 148, App
27	691.5	8.5	347	9	US-09-909-204-148	Sequence 148, App
28	691.5	8.5	347	9	US-10-230-163-114	Sequence 114, App
29	691.5	8.5	347	9	US-09-904-786-148	Sequence 148, App
30	691.5	8.5	347	9	US-09-906-646-148	Sequence 148, App
31	691.5	8.5	347	9	US-09-906-700-148	Sequence 148, App
32	691.5	8.5	347	9	US-09-902-903-148	Sequence 148, App
33	691.5	8.5	347	9	US-09-903-749A-148	Sequence 148, App
34	691.5	8.5	347	9	US-09-903-786-148	Sequence 148, App
35	691.5	8.5	347	9	US-10-218-631-114	Sequence 114, App
36	691.5	8.5	347	9	US-10-230-338-114	Sequence 114, App
37	691.5	8.5	347	9	US-09-902-736-148	Sequence 148, App
38	691.5	8.5	347	9	US-09-904-119-148	Sequence 148, App
39	691.5	8.5	347	9	US-09-904-956-148	Sequence 148, App
40	691.5	8.5	347	9	US-09-907-794-148	Sequence 148, App
41	691.5	8.5	347	9	US-10-230-414-114	Sequence 114, App
42	691.5	8.5	347	9	US-09-902-692-148	Sequence 148, App
43	691.5	8.5	347	9	US-09-903-520-148	Sequence 148, App
44	691.5	8.5	347	9	US-09-903-943-148	Sequence 148, App
45	691.5	8.5	347	9	US-09-904-462-148	Sequence 148, App

ALIGNMENTS

RESULT 1

US-10-042-431-11  
; Sequence 11, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: FRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042.431  
; PRIOR FILING DATE: 2001-10-25  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 1453  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-042-431-11

Query Match 100.0%; Score 8138; DB 9; Length 1453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MMLPQNSWHIDFGRCCHQNLFSAVVTCILLNSCLISSFNGTDLRLVNGDGPCSGT	60
Db	1	MMLPQNSWHIDFGRCCHQNLFSAVVTCILLNSCLISSFNGTDLRLVNGDGPCSGT	60
Qy	61	VEVFPQGMGTVCDDGNWNTASTVVCVKOLGCPFPFAMFRFGQAVTRHGIKILDDVSCYGN	120
Db	61	VEVFPQGMGTVCDDGNWNTASTVVCVKOLGCPFPFAMFRFGQAVTRHGIKILDDVSCYGN	120
Qy	121	ESALWECOHREWSHNCYHGVEDGVNVCYGEANGLRLVDGNNCSGRVEVFKQERWGTIC	180
Db	121	ESALWECOHREWSHNCYHGVEDGVNVCYGEANGLRLVDGNNCSGRVEVFKQERWGTIC	180

QY 181 DDGWLNTAAVVCRLGCPSPSSIFSSGVNNSPAVLPIWLDLILCOGNELALWNCRRHWG 240  
DB 181 DDGWLNTAAVVCRLGCPSPSSIFSSGVNNSPAVLPIWLDLILCOGNELALWNCRRHWG 240  
QY 241 NHDCHSHEDVTLTCYDSDLELRVLGGTNRGCMGRVELKIQGRWGTVCCHKKNNAAADVVC 300  
DB 241 NHDCHSHEDVTLTCYDSDLELRVLGGTNRGCMGRVELKIQGRWGTVCCHKKNNAAADVVC 300  
QY 301 KOLGCGTALHFAGLPHLGSGSDVWLDGVSGNSFSLWDCRHSFTVNFDCILHQNDSVI 360  
DB 301 KOLGCGTALHFAGLPHLGSGSDVWLDGVSGNSFSLWDCRHSFTVNFDCILHQNDSVI 360  
QY 361 CSDGADLELRADGNNCGSRVEVRIHQWMTICDNKWNKNEALVVKOLGCPFSVFGSR 420  
DB 361 CSDGADLELRADGNNCGSRVEVRIHQWMTICDNKWNKNEALVVKOLGCPFSVFGSR 420  
QY 421 RAKPSNEARDIWNISCTGNESALWDCYDCKAKRTCFRSDAGVICSDRADLRLVG 480  
DB 421 RAKPSNEARDIWNISCTGNESALWDCYDCKAKRTCFRSDAGVICSDRADLRLVG 480  
QY 481 AHSPCYGRLEVYQGEWGTVCCHDRWSTRNAVVCOLGCGKPMHVFGMTYFKEASGPIWL 540  
DB 481 AHSPCYGRLEVYQGEWGTVCCHDRWSTRNAVVCOLGCGKPMHVFGMTYFKEASGPIWL 540  
QY 541 DDVSCIGNESNIWDCHEHSGWGHKNCVHREDVITVCSGDTWGLRLVGGSNRCSGRLEYF 600  
DB 541 DDVSCIGNESNIWDCHEHSGWGHKNCVHREDVITVCSGDTWGLRLVGGSNRCSGRLEYF 600  
QY 601 QCRWGTVCDDGNWNSKAAAVVCQOLDCPSSIIGMGLGNASTGYKTIWLDVSCDGDSDLW 660  
DB 601 QCRWGTVCDDGNWNSKAAAVVCQOLDCPSSIIGMGLGNASTGYKTIWLDVSCDGDSDLW 660  
QY 661 SCRNCGWNNDCSHSEDEVGVICSDASDMELRLVGGSSRCAGVEVNVQAVGILCANGWG 720  
DB 661 SCRNCGWNNDCSHSEDEVGVICSDASDMELRLVGGSSRCAGVEVNVQAVGILCANGWG 720  
QY 721 MNAEVCRLGCGSAIRVSRPHETERTLHILMNSGCTGGEASLWDCIRWENKQTACH 780  
DB 721 MNAEVCRLGCGSAIRVSRPHETERTLHILMNSGCTGGEASLWDCIRWENKQTACH 780  
QY 781 LNEASLCSAHRQPLRGADMPGSRVEVKHADTWRSVCDSDFSLHAANVLCRELNGD 840  
DB 781 LNEASLCSAHRQPLRGADMPGSRVEVKHADTWRSVCDSDFSLHAANVLCRELNGD 840  
QY 841 AISLSVGDHFGKNGLTWAEKFCESSETHLALCPIVQHPEDTCHSREVGVVCSRYTIDV 900  
DB 841 AISLSVGDHFGKNGLTWAEKFCESSETHLALCPIVQHPEDTCHSREVGVVCSRYTIDV 900  
QY 901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRLSCGTALSTGGKYIGERSV 960  
DB 901 RLVNGKSQCDGQVEINVLGHWGSLCDTHWDPEDARVLCRLSCGTALSTGGKYIGERSV 960  
QY 961 RVWGHFCHLGNESLDCNQMTVLGAPPCIHGNTVSVICTGSLTOPLPCLANVSDPYLS 1020  
DB 961 RVWGHFCHLGNESLDCNQMTVLGAPPCIHGNTVSVICTGSLTOPLPCLANVSDPYLS 1020  
QY 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRIEYTHDGFWGTICDDGWDLSDAHVVCKLGGC 1080  
DB 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRIEYTHDGFWGTICDDGWDLSDAHVVCKLGGC 1080  
QY 1081 VAFNATVSAHFEGSGPIWLDLNLCTGTSHLWQCPSRGWGQHDCHRHEDAGVICSFTA 1140  
DB 1081 VAFNATVSAHFEGSGPIWLDLNLCTGTSHLWQCPSRGWGQHDCHRHEDAGVICSFTA 1140  
QY 1141 LRLYSETETESACRLEVEFYNGTWGSGVRNNTTITAGTVCRQLGCGENGVS LAPLSKT 1200  
DB 1141 LRLYSETETESACRLEVEFYNGTWGSGVRNNTTITAGTVCRQLGCGENGVS LAPLSKT 1200  
QY 1201 GSGFMVDDIQCPKTHISTWQCLSAFWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260  
DB 1201 GSGFMVDDIQCPKTHISTWQCLSAFWERRISSPAETWITCEDRIRVRGGDTECSGRVE 1260  
QY 1261 IWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALALRADSFQCGTGTIWLDDMRCKGNES 1320

DB 1261 IWHAGSWGTVCDSDWDLAEAEVVCQOLGCGSALALRADSFQCGTGTIWLDDMRCKGNES 1320  
QY 1321 FLWCHAKPWQSCDGHKEDAGVRCGQSLSKSNASSGHLALILSSIFGLLLVLFILFL 1380  
DB 1321 FLWCHAKPWQSCDGHKEDAGVRCGQSLSKSNASSGHLALILSSIFGLLLVLFILFL 1380  
QY 1381 TWCVRQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDTTPNHGCEASDT 1440  
DB 1381 TWCVRQKQKHLPLRVSTRRRGSLEENLFHEMETCLKREDPHGTRTSDTTPNHGCEASDT 1440  
QY 1441 SLLGVLPASEATK 1453  
DB 1441 SLLGVLPASEATK 1453

RESULT 2

US-09-759-130B-381  
; Sequence 381, Application US/09759130B  
; Publication No. US2003002279A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: McCarthy, Sean A  
; APPLICANT: Fraser, Christopher C  
; APPLICANT: Sharp, John D  
; APPLICANT: Barnes, Thomas S  
; APPLICANT: KIRST, Susan J  
; APPLICANT: Mackay, Charles R  
; APPLICANT: Leiby, Kevin R  
; APPLICANT: Wrighton, Nicolas  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Holtzman, Douglas A  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER  
; TITLE OF INVENTION: USES  
; FILE REFERENCE: MP100-5350MIM  
; CURRENT APPLICATION NUMBER: US/09/759,130B  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: US 09/479,249  
; PRIOR FILING DATE: 2000-01-07  
; PRIOR APPLICATION NUMBER: US 09/559,497  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/596,194  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/342,364  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: US 09/608,452  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/393,996  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: US 09/602,871  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 09/420,707  
; PRIOR FILING DATE: 1999-10-19  
; NUMBER OF SEQ ID NOS: 460  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 381  
; LENGTH: 1453  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-759-130B-381

Query Match 100.0%; Score 8138; DB 9; Length 1453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WMLPNSHIDGRCCCHONLFSVAVVTCILLNSCFLLSSFNGLTDLRLVNGDPCSGT 60  
|||||

```
Db 1 MMLPQNSWHIDFRGCCCHQNLFSAVVTCILLNSCFLISSFNCTDLELRLVNGDPCSGT 60
QY 61 VEYKFGQMGTCVDDGWNNTASTVCKQLGCPSEFAFRGQAVTRHGKTLWDDVSCYGN 120
Db 61 VEYKFGQMGTCVDDGWNNTASTVCKQLGCPSEFAFRGQAVTRHGKTLWDDVSCYGN 120
QY 121 ESALWECQHQREMGSHNCYHGEDYGVNCGYGANLGLRLVDGNNSCSGRVEYKFOERMGTC 180
Db 121 ESALWECQHQREMGSHNCYHGEDYGVNCGYGANLGLRLVDGNNSCSGRVEYKFOERMGTC 180
QY 181 DDCGNLNTAAVVCRLQGCPSSTSSGVVNSPAVLRLPWLDDIILCOGNEALWNCRRHGW 240
Db 181 DDCGNLNTAAVVCRLQGCPSSTSSGVVNSPAVLRLPWLDDIILCOGNEALWNCRRHGW 240
QY 241 NHDCSHNEVDVTLTCYDSSDLELRLVGGTNRGMGRVELKIQGRMGTCVCHHKWNAADVVC 300
Db 241 NHDCSHNEVDVTLTCYDSSDLELRLVGGTNRGMGRVELKIQGRMGTCVCHHKWNAADVVC 300
QY 301 KQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWDCRHSRGTVNFDCLRHQNDVSVI 360
Db 301 KQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWDCRHSRGTVNFDCLRHQNDVSVI 360
QY 361 CSOGADLELRLADGNNCSGRVEYRHEQWNTICDQWKNQALVYCKQLGCPSEVFGSR 420
Db 361 CSOGADLELRLADGNNCSGRVEYRHEQWNTICDQWKNQALVYCKQLGCPSEVFGSR 420
QY 421 RAKPSNEARDIWNISCTGNEALMDCTYDGRKAKRTCFRRSDAGVTCSDKADLDLRLVG 480
Db 421 RAKPSNEARDIWNISCTGNEALMDCTYDGRKAKRTCFRRSDAGVTCSDKADLDLRLVG 480
QY 481 AHSPCYGRLEVYKQGEWGTVDHWRSTRNAAVYCKQLGCGKPMHVFGMTYFKEASGPWL 540
Db 481 AHSPCYGRLEVYKQGEWGTVDHWRSTRNAAVYCKQLGCGKPMHVFGMTYFKEASGPWL 540
QY 541 DVVSCIGNESNIWDCRHSKGKNCVHREDVIYTCSDATWGLRLVGGSNRCSGRLEVYF 600
Db 541 DVVSCIGNESNIWDCRHSKGKNCVHREDVIYTCSDATWGLRLVGGSNRCSGRLEVYF 600
QY 601 QGRWGTVCDDGWSKAAAVVCSQDPCSSITIGMLGNASTGYGKIWLDDVSCDGDSDDLW 660
Db 601 QGRWGTVCDDGWSKAAAVVCSQDPCSSITIGMLGNASTGYGKIWLDDVSCDGDSDDLW 660
QY 661 SCRNWGNNDCSHSEDDVGVICSADSMELRLVGGSSRCAGKYEVNVQVAGVILCANGWG 720
Db 661 SCRNWGNNDCSHSEDDVGVICSADSMELRLVGGSSRCAGKYEVNVQVAGVILCANGWG 720
QY 721 MNTAEVVCRLQCGSALVRSRPHFTERTLHILMSNGSGCTGGEASLWDCIRWEKQTACH 780
Db 721 MNTAEVVCRLQCGSALVRSRPHFTERTLHILMSNGSGCTGGEASLWDCIRWEKQTACH 780
QY 781 LNMEASLICSAPROPRLVGMPCSGRVEYKHADTWRSVCDPSFSLHAANVLCRELNCGD 840
Db 781 LNMEASLICSAPROPRLVGMPCSGRVEYKHADTWRSVCDPSFSLHAANVLCRELNCGD 840
QY 841 AISLSVGDHFGKNGLTWAEKFCQEGSETHALCPIVQHPEDTCIHSREYGVVCSRYTDV 900
Db 841 AISLSVGDHFGKNGLTWAEKFCQEGSETHALCPIVQHPEDTCIHSREYGVVCSRYTDV 900
QY 901 RLVNGKSQCQDQVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGGKYIGERSV 960
Db 901 RLVNGKSQCQDQVEINVLGHWSGLCDTHWDPEDARVLCRLSCGTALSTTGGKYIGERSV 960
QY 961 RVNGHRPHCLGNESLNDNCQMTVLGAPPCIHGNTVSVICTGSLTOPFLPCLANVSDPYLS 1020
Db 961 RVNGHRPHCLGNESLNDNCQMTVLGAPPCIHGNTVSVICTGSLTOPFLPCLANVSDPYLS 1020
QY 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGDWLDSDAHVVCQKLGCG 1080
Db 1021 AVPEGSALICLEDKRLRLVDGDSRCAGRVEIYHDGFWGTICDDGDWLDSDAHVVCQKLGCG 1080
QY 1081 VAFNATVSAHFEGSGPWLDDLNCTGTESHLWQCPSRGWGQHDCHRHKEDAGVICSFTFA 1140
Db 1081 VAFNATVSAHFEGSGPWLDDLNCTGTESHLWQCPSRGWGQHDCHRHKEDAGVICSFTFA 1140
```

```
QY 1141 LRLYSETETESCAGRLEVFYNGTWGSGVRNNTITATAGIVCRLGGCGENGVS LAPLSKT 1200
Db 1141 LRLYSETETESCAGRLEVFYNGTWGSGVRNNTITATAGIVCRLGGCGENGVS LAPLSKT 1200
QY 1201 GSGFMWDDIQCCKTHISIWQCLSAPEWERRISSPABETWITCEDRIRVRGDTGTECSRVE 1260
Db 1201 GSGFMWDDIQCCKTHISIWQCLSAPEWERRISSPABETWITCEDRIRVRGDTGTECSRVE 1260
QY 1261 IWHAGSWGTVCCDSDWLAEEAVVCOQLGCGSALAALDASFGOGTGTIWLDDMRCKGNES 1320
Db 1261 IWHAGSWGTVCCDSDWLAEEAVVCOQLGCGSALAALDASFGOGTGTIWLDDMRCKGNES 1320
QY 1321 FLWDCHAKPMGQSDCKGHKEDAGVRCQSGLKSLNASSGHLALILSSIFGLLLVLFILFL 1380
Db 1321 FLWDCHAKPMGQSDCKGHKEDAGVRCQSGLKSLNASSGHLALILSSIFGLLLVLFILFL 1380
QY 1381 TWCVRQKQKHLPLRVSTRRRRGSLEENLFHEMETCLRKREDPHGTRTSDTNNHGCEDASDT 1440
Db 1381 TWCVRQKQKHLPLRVSTRRRRGSLEENLFHEMETCLRKREDPHGTRTSDTNNHGCEDASDT 1440
QY 1441 SLGLVLPASEATK 1453
Db 1441 SLGLVLPASEATK 1453
```

RESULT 3  
US-10-042-431-13  
; Sequence 13, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: PRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; PREVENTIVE, THERAPEUTIC, AND OTHER USES  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042,431  
; CURRENT FILING DATE: 2001-10-25  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13:  
; LENGTH: 1413  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-042-431-13

Query Match 97.2%; Score 7913; DB 9; Length 1413;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 41 FNGTDLRLVNGDPCSGTVEYKFGQMGTCVDDGWNNTASTVCKQLGCPSEFAFRF 100
Db 1 FNGTDLRLVNGDPCSGTVEYKFGQMGTCVDDGWNNTASTVCKQLGCPSEFAFRF 60
QY 101 GOAVTRHGKIWLDDVSCYGNESALWECQHQREMGSHNCYHGEDYGVNCGYGANLGLRLVDG 160
Db 61 GOAVTRHGKIWLDDVSCYGNESALWECQHQREMGSHNCYHGEDYGVNCGYGANLGLRLVDG 120
QY 161 NNSCSGRVEYKFOERMGTCICDDGNLNTAAVVCRLQGCPSSTSSGVVNSPAVLRLPWL 220
Db 121 NNSCSGRVEYKFOERMGTCICDDGNLNTAAVVCRLQGCPSSTSSGVVNSPAVLRLPWL 180
QY 221 DILCOGNELALWNCRRHGWGNDCHSHNEVDVTLTCYDSSDLELRLVGGTNRGMGRVELKIQ 280
Db 181 DILCOGNELALWNCRRHGWGNDCHSHNEVDVTLTCYDSSDLELRLVGGTNRGMGRVELKIQ 240
QY 281 GRMGTVCHHKWNAADVVCQKLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWD 340
```

Db 241 GRWGTVCCHKNNAADVVCQKCGTALHFAGLPHLQSGDVVWLDGSCNSESFLWD 300  
QY 341 CRHSCTVNFDCIHLQNDVSVICSDGADLELRLADGNNCGRVEVRIHPOWNTICDQWKN 400  
Db 301 CRHSCTVNFDCIHLQNDVSVICSDGADLELRLADGNNCGRVEVRIHPOWNTICDQWKN 360  
QY 401 EQALVVCQKLGCPFSVFGSRRAKPSNEARDIWINISICTGNESALWDCYDGGKAKRTCFR 460  
Db 361 EQALVVCQKLGCPFSVFGSRRAKPSNEARDIWINISICTGNESALWDCYDGGKAKRTCFR 420  
QY 461 RSDAGVICSADKADLRLVGAHSPCYGRLEVYKQGEWGTVCCHDRWSTRNAAVVCQKLGCG 520  
Db 421 RSDAGVICSADKADLRLVGAHSPCYGRLEVYKQGEWGTVCCHDRWSTRNAAVVCQKLGCG 480  
QY 521 KPMHVFGMTYKPEASGPTWLDVSCIGNESNIWDCHEHSGWGHKNCVHREDVIVTCSGAT 580  
Db 481 KPMHVFGMTYKPEASGPTWLDVSCIGNESNIWDCHEHSGWGHKNCVHREDVIVTCSGAT 540  
QY 581 WGLRLVGSNRCGRLEVYFQGRWGTVCDDGNWNSKAAAVVCQKLGCGNAST 640  
Db 541 WGLRLVGSNRCGRLEVYFQGRWGTVCDDGNWNSKAAAVVCQKLGCGNAST 600  
QY 641 GYKTIWLDVSCDGEDSLWSCRNSGWNNDCHSHEDVGVICSDASDMELRLVGGSSRCA 700  
Db 601 GYKTIWLDVSCDGEDSLWSCRNSGWNNDCHSHEDVGVICSDASDMELRLVGGSSRCA 660  
QY 701 GKEVNVQAGVILCANGWGNIAEVVCRQLPCGSAIRVSRPHETERTLHILMNSGCT 760  
Db 661 GKEVNVQAGVILCANGWGNIAEVVCRQLPCGSAIRVSRPHETERTLHILMNSGCT 720  
QY 761 GGEASLWDCIRWEKQTAACHLWNEASLICSRAHQRLVGMADPCSGRVEVHADTWRSVC 820  
Db 721 GGEASLWDCIRWEKQTAACHLWNEASLICSRAHQRLVGMADPCSGRVEVHADTWRSVC 780  
QY 821 DSDFSLSHAANVLCRELNCDAISLVGDHFGKNGLTWAEKFCQEGSETHALCPVQHP 880  
Db 781 DSDFSLSHAANVLCRELNCDAISLVGDHFGKNGLTWAEKFCQEGSETHALCPVQHP 840  
QY 881 EDTCHSRREVGVVCSRYTDLVRLVNGKSCDGOVEINVLGHWSLCTDTHWDPEDARVLCRQ 940  
Db 841 EDTCHSRREVGVVCSRYTDLVRLVNGKSCDGOVEINVLGHWSLCTDTHWDPEDARVLCRQ 900  
QY 941 LSCGTALSTGGKYIGERSVRVWGHFHCILGNESLLDNCQMTVLGAPPCIHGNTVSVICT 1000  
Db 901 LSCGTALSTGGKYIGERSVRVWGHFHCILGNESLLDNCQMTVLGAPPCIHGNTVSVICT 960  
QY 1001 GSLTQPLFCLANVSDPYLSAVPEGSALICLEDKRLRLVGDSCRAGRVEIYHDFGWTI 1060  
Db 961 GSLTQPLFCLANVSDPYLSAVPEGSALICLEDKRLRLVGDSCRAGRVEIYHDFGWTI 1020  
QY 1061 CDGWDLSDAHVVYVQKLGCGVAFNATVSAHFEGGSGPIWLDLNCCTGESHVWQCPSRGW 1120  
Db 1021 CDGWDLSDAHVVYVQKLGCGVAFNATVSAHFEGGSGPIWLDLNCCTGESHVWQCPSRGW 1080  
QY 1121 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVYNGTWGSGRRNITTAIGIV 1180  
Db 1081 GQHDCRHKEDAGVICSEFTALRLYSETETESCAGRLEVYNGTWGSGRRNITTAIGIV 1140  
QY 1181 CRQLCGGNGVVS LAPLSKTGSGFMWVDIQC PKTHISTWQCLSAFWERRISSPAEETWI 1240  
Db 1141 CRQLCGGNGVVS LAPLSKTGSGFMWVDIQC PKTHISTWQCLSAFWERRISSPAEETWI 1200  
QY 1241 TCEDRLVRGGTDECSGRVEIWHAGSWGTVCCDSDWDLAEAVVYVQKLGCGSALALRDAS 1300  
Db 1201 TCEDRLVRGGTDECSGRVEIWHAGSWGTVCCDSDWDLAEAVVYVQKLGCGSALALRDAS 1260  
QY 1301 FQGTGCTIWLDDMRCKGNESFVLWDCHAKPWQSDCGCHKEDAGVRCGGQSLKSLNASSGHL 1360  
Db 1261 FQGTGCTIWLDDMRCKGNESFVLWDCHAKPWQSDCGCHKEDAGVRCGGQSLKSLNASSGHL 1320  
QY 1361 ALILSSIFGLLILVLFILFTWCRVQKOKHPLRVSTRRGSLEENLFHEMETCLKREDP 1420

Db 1321 ALILSSIFGLLILVLFILFTWCRVQKOKHPLRVSTRRGSLEENLFHEMETCLKREDP 1380  
QY 1421 HGRTRSDTTPNHGCEADSDTSLGVLPAASEATK 1453  
Db 1381 HGRTRSDTTPNHGCEADSDTSLGVLPAASEATK 1413  
RESULT 4  
US-09-759-130B-383  
; Sequence 383, Application US/09759130B  
; Publication No. US20030022279A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: McCarthy, Sean A  
; APPLICANT: Fraser, Christopher C  
; APPLICANT: Sharp, John D  
; APPLICANT: Barnes, Thomas S  
; APPLICANT: KIRST, Susan J  
; APPLICANT: Mackay, Charles R  
; APPLICANT: Myers, Paul S  
; APPLICANT: Leiby, Kevin R  
; APPLICANT: Wrighton, Nicolas  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Holtzman, Douglas A  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER  
; TITLE OF INVENTION: USES  
; FILE REFERENCE: MP100-5350NM1  
; CURRENT APPLICATION NUMBER: US/09759,130B  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: US 09/479,249  
; PRIOR FILING DATE: 2000-01-07  
; PRIOR APPLICATION NUMBER: US 09/559,497  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/596,194  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/342,364  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: US 09/608,452  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/393,996  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: US 09/602,871  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 09/420,707  
; NUMBER OF SEQ ID NOS: 460  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 383  
; LENGTH: 1413  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-759-130B-383  
Query Match 97.2%; Score 7913; DB 9; Length 1413;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 41 FNGTDLRLVNGDPCSGTVEVKFGQGWGTVCDDGNWNTTASTVVCQKLGCGFSFAMFRF 100  
Db 1 FNGTDLRLVNGDPCSGTVEVKFGQGWGTVCDDGNWNTTASTVVCQKLGCGFSFAMFRF 60  
QY 101 GOAVTRHGKIWLDDVSCYGNESALWECOHREWSHNCYHGEDVGVNCEANGLRLVDG 160  
Db 61 GOAVTRHGKIWLDDVSCYGNESALWECOHREWSHNCYHGEDVGVNCEANGLRLVDG 120  
QY 161 NNSCSGRVEVKFOERWGTICDDGNWNLNTAAVVCQKLGCGFSFISGVNNSPAVLRLPWL 220  
Db 121 NNSCSGRVEVKFOERWGTICDDGNWNLNTAAVVCQKLGCGFSFISGVNNSPAVLRLPWL 180

QY 221 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTCTYDSDLELRLVGGTNRNCRMGVELKIQ 280  
DB 181 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTCTYDSDLELRLVGGTNRNCRMGVELKIQ 240  
QY 281 GRGVTCHHKNNAADAVCKQLGCGTALHFAGLPHLQSGSDVVMWLVGVSCSGNESFLWD 340  
DB 241 GRGVTCHHKNNAADAVCKQLGCGTALHFAGLPHLQSGSDVVMWLVGVSCSGNESFLWD 300  
QY 341 CRSGTWNFCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 400  
DB 301 CRSGTWNFCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 360  
QY 401 EQALVVKQLGCGPFSVFGSRRAKPSNEARDIWNISICTGNESALWDCYDGRKAKRTCFR 460  
DB 361 EQALVVKQLGCGPFSVFGSRRAKPSNEARDIWNISICTGNESALWDCYDGRKAKRTCFR 420  
QY 461 RSDAGVTCSDKADLDRLVGAHSPCYGRLEVYQGEWGTVCHDRWSTRNAVVCVKQLGCG 520  
DB 421 RSDAGVTCSDKADLDRLVGAHSPCYGRLEVYQGEWGTVCHDRWSTRNAVVCVKQLGCG 480  
QY 521 KPMHVFGMTYFKASGPIWLDVSVICSGNESINWDCSHSGKNCVHREDVIVTCSGDAT 580  
DB 481 KPMHVFGMTYFKASGPIWLDVSVICSGNESINWDCSHSGKNCVHREDVIVTCSGDAT 540  
QY 581 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNKAAAVVCSQLDCPSSIIIGMGLGNAST 640  
DB 541 WGLRLVGGSNRCSGRLEVYFQGRWGTVCDDGWNKAAAVVCSQLDCPSSIIIGMGLGNAST 600  
QY 641 GYCKIWLDDVSCDGDSDLSNCRSGNNDCHSDSHEDVGVICSDASDMLRLVGGSSRCA 700  
DB 601 GYCKIWLDDVSCDGDSDLSNCRSGNNDCHSDSHEDVGVICSDASDMLRLVGGSSRCA 660  
QY 701 GKVEVNVQAGVILCANGWGMNTAEVVCROLCEGSAIRVSRPHFTERTLHILMNSGCT 760  
DB 661 GKVEVNVQAGVILCANGWGMNTAEVVCROLCEGSAIRVSRPHFTERTLHILMNSGCT 720  
QY 761 GGEASLWDCIRWEWKOTACHLNMEASLCSAHRQPRLVGADMPSCGRVEVKHADTWRSVC 820  
DB 721 GGEASLWDCIRWEWKOTACHLNMEASLCSAHRQPRLVGADMPSCGRVEVKHADTWRSVC 780  
QY 821 DSDFSLHAANVLCRELNGDAISLUSVDHFGKGNGLTWAEKFCQEGSETHALALCPVQHP 880  
DB 781 DSDFSLHAANVLCRELNGDAISLUSVDHFGKGNGLTWAEKFCQEGSETHALALCPVQHP 840  
QY 881 EDTCIHSREVGVCSTRYTDVRLVNGKSCQDGOVEINVLGHWSLUCDTHWDPEDARVLCRQ 940  
DB 841 EDTCIHSREVGVCSTRYTDVRLVNGKSCQDGOVEINVLGHWSLUCDTHWDPEDARVLCRQ 900  
QY 941 LSCGTALSTTGKYGIGRSVRVMGHRPHCLGNESLDDNCOMVTVLGAPPCIHGNTVSVCIT 1000  
DB 901 LSCGTALSTTGKYGIGRSVRVMGHRPHCLGNESLDDNCOMVTVLGAPPCIHGNTVSVCIT 960  
QY 1001 GSLTQPLFPCLANVSDPYLSAVPESGALICLEDKRLRLVDGDSRCAGRVEIYHDFGWTI 1060  
DB 961 GSLTQPLFPCLANVSDPYLSAVPESGALICLEDKRLRLVDGDSRCAGRVEIYHDFGWTI 1020  
QY 1061 CDDGWLSDARHVVCKLGGCVAFNATVSAHFEGSGPIWLDLNCCTGESHLMWQPSRGW 1120  
DB 1021 CDDGWLSDARHVVCKLGGCVAFNATVSAHFEGSGPIWLDLNCCTGESHLMWQPSRGW 1080  
QY 1121 GQHDCRHKEDAGVTCSEFTALRLYSETETESACRLEVFYNGTWGSGVRRNITTAIGV 1180  
DB 1081 GQHDCRHKEDAGVTCSEFTALRLYSETETESACRLEVFYNGTWGSGVRRNITTAIGV 1140  
QY 1181 CRQLGCCGNGVSLAPLSKTSKSGFPMWDDIQCPTKTHISINWQLSAPHERLSSPAETWI 1240  
DB 1141 CRQLGCCGNGVSLAPLSKTSKSGFPMWDDIQCPTKTHISINWQLSAPHERLSSPAETWI 1200  
QY 1241 TCEDRIIRVRGDTPCSGRVEIWHAGSNGTVCDDSDMLAEAEVVCQOLGCCSALAALRDAS 1300  
DB 1201 TCEDRIIRVRGDTPCSGRVEIWHAGSNGTVCDDSDMLAEAEVVCQOLGCCSALAALRDAS 1260

QY 1301 FGOGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSNASSGHL 1360  
DB 1261 FGOGTGTIWLDDMRCKGNESFLWDCHAKPWGQSDCGHKEDAGVRCSGQSLKSNASSGHL 1320  
QY 1361 ALILSSIFGGLLLVFLFLTWCRVQKOKHLPLRYSTRRRGSLLENLFHEMETCLKREDP 1420  
DB 1321 ALILSSIFGGLLLVFLFLTWCRVQKOKHLPLRYSTRRRGSLLENLFHEMETCLKREDP 1380  
QY 1421 HGRTSDTTPNHGCEADSDTSLIGVLPAASEATK 1453  
DB 1381 HGRTSDTTPNHGCEADSDTSLIGVLPAASEATK 1413

RESULT 5  
US-10-042-431-14  
; Sequence 14, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: BRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042,431  
; CURRENT FILING DATE: 2001-10-25  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1319  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-042-431-14

Query Match 91.3%; Score 7426; DB 9; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 FNGTDLRLVNGDPCSGTVEVKFQGWGTVCDDGWNNTASTVVCVKQLGCPFSFAMPRF 100  
DB 1 FNGTDLRLVNGDPCSGTVEVKFQGWGTVCDDGWNNTASTVVCVKQLGCPFSFAMPRF 60  
QY 101 GOAVTRHGKIWLDDVSVICGNESALWECOHREWGSHNCHYHGEDVGVNCGYGEANLGLRLVDG 160  
DB 61 GOAVTRHGKIWLDDVSVICGNESALWECOHREWGSHNCHYHGEDVGVNCGYGEANLGLRLVDG 120  
QY 161 NNSCGRVEVKFQBRWGTICDDGWNLTAAVVCROLGCPSSFISGGVNSPAVLRLPIWLD 220  
DB 121 NNSCGRVEVKFQBRWGTICDDGWNLTAAVVCROLGCPSSFISGGVNSPAVLRLPIWLD 180  
QY 221 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTCTYDSDLELRLVGGTNRNCRMGVELKIQ 280  
DB 181 DILCOGNELALWNCRRHGWGNDHCSHNEEDVLTCTYDSDLELRLVGGTNRNCRMGVELKIQ 240  
QY 281 GRGVTCHHKNNAADAVCKQLGCGTALHFAGLPHLQSGSDVVMWLVGVSCSGNESFLWD 340  
DB 241 GRGVTCHHKNNAADAVCKQLGCGTALHFAGLPHLQSGSDVVMWLVGVSCSGNESFLWD 300  
QY 341 CRSGTWNFCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 400  
DB 301 CRSGTWNFCLHONDVSVICSGADLELRLADGNNCSGRVEVRIHEQWNTICDQWKN 360  
QY 401 EQALVVKQLGCGPFSVFGSRRAKPSNEARDIWNISICTGNESALWDCYDGRKAKRTCFR 460  
DB 361 EQALVVKQLGCGPFSVFGSRRAKPSNEARDIWNISICTGNESALWDCYDGRKAKRTCFR 420  
QY 461 RSDAGVTCSDKADLDRLVGAHSPCYGRLEVYQGEWGTVCHDRWSTRNAVVCVKQLGCG 520

Db 421 RSDAGVICSDKADLDLRLVLGHAHSPCYGRLEVKYQGEWGTVCCHDRWSTRNAAVVCKQLGCG 480  
Qy 521 KPMHVFMTYFKEASGPIWLDVDSICGNESINWDEHSGWGHKNCVHREDVIVTCSGDAT 580  
Db 481 KPMHVFMTYFKEASGPIWLDVDSICGNESINWDEHSGWGHKNCVHREDVIVTCSGDAT 540  
Qy 581 WGLRLVGSNRCGRLEVYFQGRWGTVCDDGWNKAAAVVCSQDLPSSITGMGLGNAST 640  
Db 541 WGLRLVGSNRCGRLEVYFQGRWGTVCDDGWNKAAAVVCSQDLPSSITGMGLGNAST 600  
Qy 641 GYKLTWLDVSCDGDSDLSNCRSGWGNDCSHSDVGVICSDASDMLRLVGSRSRA 700  
Db 601 GYKLTWLDVSCDGDSDLSNCRSGWGNDCSHSDVGVICSDASDMLRLVGSRSRA 660  
Qy 701 GKVEVVOGAVGILCANGWGNIAEVVCRQLECGSAIRVSREPHTERTLHLMSNSGCT 760  
Db 661 GKVEVVOGAVGILCANGWGNIAEVVCRQLECGSAIRVSREPHTERTLHLMSNSGCT 720  
Qy 761 GGEASLWPCIRWVKOTACHLNMEASLICSARHROPRLVGADMPGSGRVEVVKHDTWRSVC 820  
Db 721 GGEASLWPCIRWVKOTACHLNMEASLICSARHROPRLVGADMPGSGRVEVVKHDTWRSVC 780  
Qy 821 DSDFSLHAANVLCRELNGCDALISVGDHFGKNGGLTWAERFQCEGSETHLALCPVIOHP 880  
Db 781 DSDFSLHAANVLCRELNGCDALISVGDHFGKNGGLTWAERFQCEGSETHLALCPVIOHP 840  
Qy 881 EDTCTHSREVGVVCSRYTDRVLVNGKSCDGOVEINVLGHWSGLCDTHWDPEDARVLCRQ 940  
Db 841 EDTCTHSREVGVVCSRYTDRVLVNGKSCDGOVEINVLGHWSGLCDTHWDPEDARVLCRQ 900  
Qy 941 LSCGTALSTTGKYGISRVVRWGHRRFCLGNESLDDNCOMTVLGPACIIGNTVSVICT 1000  
Db 901 LSCGTALSTTGKYGISRVVRWGHRRFCLGNESLDDNCOMTVLGPACIIGNTVSVICT 960  
Qy 1001 GSLTOPLPFCLANVSDPYLSAPEGSAICLEDKRLRLVDGDSRCAGRVEVYHDFGWTI 1060  
Db 961 GSLTOPLPFCLANVSDPYLSAPEGSAICLEDKRLRLVDGDSRCAGRVEVYHDFGWTI 1020  
Qy 1061 CDGWDLSDAHVVCKQLGCGVAFNATVSAHFEGSGPIWLDLNCGTGESHLMQCPSRGW 1120  
Db 1021 CDGWDLSDAHVVCKQLGCGVAFNATVSAHFEGSGPIWLDLNCGTGESHLMQCPSRGW 1080  
Qy 1121 GOHDCRHKEDAGVICSERTALRYSETETESCAGRLEVFYNGTWGSGRRNITTAIGIV 1180  
Db 1081 GOHDCRHKEDAGVICSERTALRYSETETESCAGRLEVFYNGTWGSGRRNITTAIGIV 1140  
Qy 1181 CRQLGCGGVVSLAPLSKTSKGGFMWDDIQCPKTHSIWQCLSAPEWERRISSPAEETWI 1240  
Db 1141 CRQLGCGGVVSLAPLSKTSKGGFMWDDIQCPKTHSIWQCLSAPEWERRISSPAEETWI 1200  
Qy 1241 TCEDRIRVRGGTEGSGRVEIWHAGSWGTVCDSDWDLAEAEVVCQQLGCGSALALRDAS 1300  
Db 1201 TCEDRIRVRGGTEGSGRVEIWHAGSWGTVCDSDWDLAEAEVVCQQLGCGSALALRDAS 1260  
Qy 1301 FCGGTGTWLDMDRCKGNESFLWDCCHAKPWGSDCGHREDAGVRCGSLKSLNASSGH 1359  
Db 1261 FCGGTGTWLDMDRCKGNESFLWDCCHAKPWGSDCGHREDAGVRCGSLKSLNASSGH 1319

## RESULT 6

US-09-759-130b-384

; Sequence 384, Application US/09759130B

; Publication No. US2003002279A1

; GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: McCarthy, Sean A

; APPLICANT: Fraser, Christopher C

; APPLICANT: Sharp, John D

; APPLICANT: Barnes, Thomas S

; APPLICANT: Kirt, Susan J

; APPLICANT: Mackay, Charles R

; APPLICANT: Myers, Paul S

; APPLICANT: Letby, Kevin R

; APPLICANT: Wrighton, Nicolas  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Holtzman, Douglas A  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER  
; TITLE OF INVENTION: USES  
; FILE REFERENCE: MF100-5350MIM  
; CURRENT APPLICATION NUMBER: US/09/759,130B  
; CURRENT FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: US 09/479,249  
; PRIOR FILING DATE: 2000-01-07  
; PRIOR APPLICATION NUMBER: US 09/559,497  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/596,194  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/342,364  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: US 09/608,452  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/393,996  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: US 09/602,871  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 09/420,707  
; PRIOR FILING DATE: 1999-10-19  
; NUMBER OF SEQ ID NOS: 460  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 384  
; LENGTH: 1319  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-759-130b-384

Query Match 91.3%; Score 7426; DB 9; Length 1319;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 FNGTDLRLVNGDPCSGTVEVKFQGWGTVCDDGWNNTASTVVCCKQLGCPFSFAMPRF 100  
Db 1 FNGTDLRLVNGDPCSGTVEVKFQGWGTVCDDGWNNTASTVVCCKQLGCPFSFAMPRF 60  
Qy 101 GOAVTRHGKIWLDDVSCYGNESALMECQHRWGSNCHYHGEDVGVNCGEANLGLRLVDG 160  
Db 61 GOAVTRHGKIWLDDVSCYGNESALMECQHRWGSNCHYHGEDVGVNCGEANLGLRLVDG 120  
Qy 161 NNCSGRVEVKFOERWGTICDDGWNLTAAVVCROLGCPSSFISSGVVNSPAVLRLPWL 220  
Db 121 NNCSGRVEVKFOERWGTICDDGWNLTAAVVCROLGCPSSFISSGVVNSPAVLRLPWL 180  
Qy 221 DILCOGNELALWNCRRHGRGNHDCSHNEDVTITCYDSSDLELRLVGGTNRMCGRVELKIQ 280  
Db 181 DILCOGNELALWNCRRHGRGNHDCSHNEDVTITCYDSSDLELRLVGGTNRMCGRVELKIQ 240  
Qy 281 GRWGTVCCHKNNAADVVCKQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWD 340  
Db 241 GRWGTVCCHKNNAADVVCKQLGCGTALHFAGLPHLQSGSDVWLDGVSCSGNESFLWD 300  
Qy 341 CRHSGTVNFDCLHONDVSVICSDGADLELRLADGNNCSGRVEVRHIEQWWTICDQNWKN 400  
Db 301 CRHSGTVNFDCLHONDVSVICSDGADLELRLADGNNCSGRVEVRHIEQWWTICDQNWKN 360  
Qy 401 EQALVVCKQLGCPFSVFGSRRRAKPSNEARDIWINISICTGNESALWDCTYDCAKRTCFR 460  
Db 361 EQALVVCKQLGCPFSVFGSRRRAKPSNEARDIWINISICTGNESALWDCTYDCAKRTCFR 420  
Qy 461 RSDAGVICSDKADLDLRLVLGHAHSPCYGRLEVKYQGEWGTVCCHDRWSTRNAAVVCKQLGCG 520  
Db 421 RSDAGVICSDKADLDLRLVLGHAHSPCYGRLEVKYQGEWGTVCCHDRWSTRNAAVVCKQLGCG 480

Qy 521 KPMHVFMTYKEASGPIWLDVSCIGNESINWDCSHGKNCVHREDVIVTCSDGAT 580  
Db 481 KPMHVFMTYKEASGPIWLDVSCIGNESINWDCSHGKNCVHREDVIVTCSDGAT 540  
Qy 581 WGLRLVGSNRCGRLEVFYFQGRNGTVCDDGWSKAAAVVCSQDPCSSIIIGMLGNAST 640  
Db 541 WGLRLVGSNRCGRLEVFYFQGRNGTVCDDGWSKAAAVVCSQDPCSSIIIGMLGNAST 600  
Qy 641 GYGIWLDVSCDGDSDSLNCRSGNNDCHSDSHEDVGVICSDASDMELRLVGSRRCA 700  
Db 601 GYGIWLDVSCDGDSDSLNCRSGNNDCHSDSHEDVGVICSDASDMELRLVGSRRCA 660  
Qy 701 GKVEVVOGAVGILCANGWGNIAEVVCRQLCCGSAIRVSRPHFTERTLHILMSNSGCT 760  
Db 661 GKVEVVOGAVGILCANGWGNIAEVVCRQLCCGSAIRVSRPHFTERTLHILMSNSGCT 720  
Qy 761 GGEASLWDCIRWEKOTACHLNMEASLICSARHQPRLVGADMPGSRVVEVXHADTWRSVC 820  
Db 721 GGEASLWDCIRWEKOTACHLNMEASLICSARHQPRLVGADMPGSRVVEVXHADTWRSVC 780  
Qy 821 DSDFSLHAANYLCLRELNCGDAISLSVGDHFGKGNGLTWAERFQCEGSETHLALCPVIOHP 880  
Db 781 DSDFSLHAANYLCLRELNCGDAISLSVGDHFGKGNGLTWAERFQCEGSETHLALCPVIOHP 840  
Qy 881 EDTCHSREVGVVCSRYTDLRLVNGKSQCDCQVEINVLGHWSGLCDTHWDPEDARVLCRQ 940  
Db 841 EDTCHSREVGVVCSRYTDLRLVNGKSQCDCQVEINVLGHWSGLCDTHWDPEDARVLCRQ 900  
Qy 941 LSCGTALSTTGKYGIGERSVRWGHRRFHLCLGNESLLDNCOMTVLGAPPCIHGNTVSVICT 1000  
Db 901 LSCGTALSTTGKYGIGERSVRWGHRRFHLCLGNESLLDNCOMTVLGAPPCIHGNTVSVICT 960  
Qy 1001 GSLTQPLFPCLANVSDPYLSAVPEGSAICLEDKRLRLVDGDSRCAGRVEIYHDFGWTGI 1060  
Db 961 GSLTQPLFPCLANVSDPYLSAVPEGSAICLEDKRLRLVDGDSRCAGRVEIYHDFGWTGI 1020  
Qy 1061 CDDGWDLSDAHVCOKLGGCVAFNATVSAHFGEGSGPIWLDLNCGTGESHLMWCPSRGW 1120  
Db 1021 CDDGWDLSDAHVCOKLGGCVAFNATVSAHFGEGSGPIWLDLNCGTGESHLMWCPSRGW 1080  
Qy 1121 GOHDCRHKEDAGVICSEPTALRLYSETETESCAGRLEVFYNGTWGSGVGRNITTAIAGIV 1180  
Db 1081 GOHDCRHKEDAGVICSEPTALRLYSETETESCAGRLEVFYNGTWGSGVGRNITTAIAGIV 1140  
Qy 1181 CRQLCCGNGVYVSLAPLSKTSKSGFWVDDIOCPKTHISIWOCLSAPWERRISSPAETWI 1240  
Db 1141 CRQLCCGNGVYVSLAPLSKTSKSGFWVDDIOCPKTHISIWOCLSAPWERRISSPAETWI 1200  
Qy 1241 TCEDRIRVRGGDTESGRVEIWHAGSWGTVCCDSDSLAEAEVVCQQLCGSALAALRDAS 1300  
Db 1201 TCEDRIRVRGGDTESGRVEIWHAGSWGTVCCDSDSLAEAEVVCQQLCGSALAALRDAS 1260  
Qy 1301 FGQGTGTTIWLDDMRCKGNESFLWDCCHAKPWGQSGCGHEDAGVRCGSGSLKSLNASSGH 1359  
Db 1261 FGQGTGTTIWLDDMRCKGNESFLWDCCHAKPWGQSGCGHEDAGVRCGSGSLKSLNASSGH 1319

## RESULT 7

US-09-977-577-13

; Sequence 13, Application US/09977577

; Patent No. US20020155995A1

; GENERAL INFORMATION:

; APPLICANT: MOESTRUP, Soren

; APPLICANT: MOLLER, Holger J.

; FILE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USES T

; CURRENT APPLICATION NUMBER: US/09/977.577

; CURRENT FILING DATE: 2001-10-16

; PRIOR APPLICATION NUMBER: US 60/270,120

; PRIOR FILING DATE: 2001-02-22

; PRIOR APPLICATION NUMBER: DK PA 2001 00039

; PRIOR FILING DATE: 2001-01-11

; PRIOR APPLICATION NUMBER: DK PA 2000 01543

; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 1151  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-577-13

Query Match 42.3%; Score 3441.5; DB 9; Length 1151;  
Best Local Similarity 55.2%; Pred. No. 9e-239;  
Matches 608; Conservative 166; Mismatches 311; Indels 17; Gaps 6;

Qy 364 GADLELRLLADSNCSGRVEVRIIHEQWMTICDQNNKQALVYVQLGCPESVFSGRRAK 423  
Db 41 GTDKELRLVDGENKCSGRVEVKVQEGWGTVCNNGWSMEASVVICNLQCPATAKPGWAN 100  
Qy 424 PSNEARDIWINISICTGNESALWDCITYDGKAKRT-CFRRSADAGVTCSDKADLDLRLVGAH 482  
Db 101 SSAGSGRIWMDHVSRCGNESALWDCIKHDGKHSNCTHQDDAGVTCSDGSLNEMRLTRGG 160  
Qy 483 SPCYGRLEVKYQGGMTVCHDRWSTRNAAVVCKQLGCGKPMHVFMTYFKEASGPIWLD 542  
Db 161 NWCGRIRIETKQGRWGTVCDDNFNDHASVTCRQLECSAVSFSGSSNFGEGSGPIWFD 220  
Qy 543 VSCIGNESINWDCSHGKNCVHREDVIVTCSDATMGLRLVGVGSRNCSGRLEVFYFQ 602  
Db 221 LICNGESALWNCIKHOGKNCNCHADAGVICSAGADLSRLVDGVTECSGRLEVRFOG 280  
Qy 603 RWGTVDDGWSKAAAVVCSQDPCSSIIIGMLGNASTGYGKIWLDDVSCDGDSDSLNASC 662  
Db 281 EWGTTCDGWDSDYDAVACQKLGCPATAIGRVNASKGFGHIWLDSDVSCQGHFPAVWQC 340  
Qy 663 ENSGNGNDCSHSDGVVICSASDMELRLVGVGSRNCSGRVNVQGVGILCANGWGN 722  
Db 341 KHHEWKYCHNHNEDAGVTCSDGSDLELRLRGGRGACGTVEVEIQRLLGVKCYDRGWLK 400  
Qy 723 IAEVVCRLQEGCSAIRVSRPHFTERTLHILMSNGCTGGEASLWDCIRWEKOTACHLN 782  
Db 401 EADVVCRLQEGCSAIRVSRPHFTERTLHILMSNGCTGGEASLWDCIRWEKOTACHLN 740  
Qy 783 NEASLICSARHQPRLVGADMPGSRVVEVXHADTWRSVCDSDSLHAANYLCLRELNCGDAI 842  
Db 461 EEAKTCSAHREPRVLGGDIPCSGRVEVKGHTWGSICDSDSLHAANYLCLRELNCGTVV 520  
Qy 843 SLSVGDHFGKGNLTWAEKFCQEGSETHLALCPVIOHPEDTCHSREVGVVCSRYTDLRL 902  
Db 521 SLTGAHFGEGNGQIWAEEFQCEGHEHSLSLCPAPRPGTCSHSDRVGVVCSRYTEIRL 580  
Qy 903 YNGKSQCQDQVEINVLGHWSGLCDTHWDPEDARVLCRQLSCGTALSTTGKYGIGERSVRV 962  
Db 581 YNGKTPCEGRVELKTLGANGSLCNSHWDIEDAHLVLCQQLKCGVALSTPGAREFGKNGOI 640  
Qy 963 WGRHFCILGNESLLDNCOMTVLGAPPCIHGNTVSVICTGSLTQPLFPCLANVSDPYLSAV 1022  
Db 641 WRHMFCTGTGTOHMGDCPVTALGALCPSQEVASVICSNGSQSLSSSSSLSGTPRTPI 700  
Qy 1023 PEGSALICLEDKRLRLVDGDSRCAGRVEIYHDFGWTICDDGWDLSDAHVCOKLGGCV 1082  
Db 701 PEESAVACIESGQLRLVNGGRCAGRVEIYHFGSGWTICDDSDWLDSDAHVVCRLQCGGA 760  
Qy 1083 FNATVSAHFGEGSGPIWLDLNCGTGESHLMWCPSRGWQHDCRHKEDAGVICSEPTALR 1142  
Db 761 INATOSAHFGEGETGPIWLDLNCGTGESHLMWCPSRGWQHDCRHKEDAGVICSEPTALR 820  
Qy 1143 LYSETETESCAGRLEVFYNGTWGSGVGRNITTAIAGIVCRQLCCGNGVYVSLAPLSKTS 1202  
Db 821 LTSEASREACAGRLEVFYNGTWGSGVGRNITTAIAGIVCRQLCCGNGVYVSLAPLSKTS 880  
Qy 1203 CFMWVDDIQCPTKTHISIWOCLSAPWERRISSPAETWITCEDRIRVRGGDTESGRVEI 1262  
Db 881 IPMWYDNYOCPKGPDTLWQCPSPWKEKRLASPEETWITCDNKRILQEGPTSCSGRVEIW 940

QY 1263 HAGSWGTCDDSDWDLAEAEVWCQQLGCGSALAALRDASFGQGTGTWLDMDRCCKNESFL 1322  
DB 941 HGGSWGTCDDSDWDLDAQVVCQQLGCGPALKAFKAEAFEGQGTGTWLDNEVKCKNESL 1000  
QY 1323 WCHAKPWGSCGCKHEDAGVRCQSGLK-----SLNASSGHLALILSSIFGLLLVLF 1376  
DB 1001 WDCPARRWGHSECGHEDAAVNCTDISVQKTPQKATGRSSQSSFIAGVILGVVLLAIF 1060  
QY 1377 I--LFTWCRVQOKHPLRVSTRRSGSLEENLFHEMETCKLKREDPHGTRTSDDDTPNHGC 1434  
DB 1061 VALFELTKRRQROR---LAVSSRGENLVHQIYQREMNNSCLNADDLDMNSSEN--SHES 1115  
QY 1435 EDASDTSILGV---LPASATK 1453  
DB 1116 ADFAAELISVSKFLPISGMK 1137

## RESULT 8

US-09-977-577-10  
; Sequence 10, Application US/09977577  
; Patent No. US2002015595A1  
; GENERAL INFORMATION:  
; APPLICANT: MOESTRUP, Soren  
; APPLICANT: MOLLER, Holger J.  
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USES T  
; FILE REFERENCE: MOESTRUP-1A  
; CURRENT APPLICATION NUMBER: US/09/977,577  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: US 60/270,120  
; PRIOR FILING DATE: 2001-02-22  
; PRIOR APPLICATION NUMBER: DK PA 2001 00039  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: DK PA 2000 01543  
; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 1116  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-577-10

Query Match 42.1%; Score 3429; DB 9; Length 1116;  
Best Local Similarity 56.2%; Pred. No. 6.8e-238;  
Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

QY 364 GADLELRADGNNCGSRVEVRIHEOWTTCIDQNNKNEQALVVCQKQCPFSVFGSRRAK 423  
DB 41 GTDKELRLVDGNNCKSGRVEVVKVQEWGTVCNNGWSMEAVSVICQKQCPAIAKAPGAN 100  
QY 424 PSNEARDIWNISCTGNESALWDCTYDGAKRT-CFRSDAGVICSADKADLRLVGAH 482  
DB 101 SSAGSRIWMDHVSRCGNESALWDCKDHWGKHSNTHQDAGVTCSDGNSLEMLRTRG 160  
QY 483 SPYGRLEVKYOGWGTCHDRWSTRNAVVCQKQCGPMMVFMGTYPKEASGPTWLD 542  
DB 161 NMCSGRIEIKFQWGTGTCDDNFNDHSAVICQKQCGSAVSFGSSNFGSGPTWFD 220  
QY 543 VSCIGNESINWDCSHSGWKNCVHREDVITVCSGDTWGLRLVGSNRCGRLEVVYFQ 602  
DB 221 LICNGNESALWCKHOGWGNKNDHAEDAGVICSADLSRLVDGVTCSGRLEVRFG 280  
QY 603 RMGTVCDDGWNKAAAVVCQQLGCGPSSIIICMGLGNASTGYGKIWLDDVSCDGDDES 662  
DB 281 EMGTICDDGWDSDYDAVACKQCGPATAVTAIGRVNASKGFGHILWDSVSCQGHAPVWC 340  
QY 663 RNSGWNDCSHSDVGVVTCSDASDMELRVGSSRCAGKVEVNVQGVGILCANGWGN 722  
DB 341 KHHEWGHKVCNHNEDAGVTCSDGSDLELRGSGRCAGTVEVIGRLGKVCDCRGWGLK 400  
QY 723 IAEVVCQKQCGSALIRVSRPHPTERTLHILMSNCGTGGESALWDCIRWENKQTAACHLN 782  
DB 401 EADVVCQKQCGSALATYSQVYSKIQATWLVFLSSCNGNETSLWDCKNQWGGGLTCDHY 460

QY 783 MEASLICSARQPRILVGDMPGCSGRVEVVKHADTWRSVCDSDSFLHAANYLCRELNGDAI 842  
DB 461 BEAKITCSARHREPRILVGGDIPCSGRVEVVKHGTWTGSCICSDSFLAASVLCRELQCTTV 520  
QY 843 SLVSVDHFGKGNLWAEKFCQEGSETHALCPIVQHPEDTCTIHSREVGVVCSRTIDVRL 902  
DB 521 SLTGAHFGEGNGQIWAEEFQCEGHESHLCLCPVPRPEGTCSHSDVGVVCSRTYFIRL 580  
QY 903 VNGKSOCDQOEVINVLGHWSLCLDTHWDPEDARVLCRQLSCGTALSTTGGKVIERSVVR 962  
DB 581 VNGKTPCGRVELKTLGAWGSLCNHSHWIDEDAHVLCQKLCGVALSTPGARFGRNGQI 640  
QY 963 WGRFHCLGNESLNDNCOMTVLGAAPPCIHGNTVSVICTSLTOPPLPCIANVSDPLVSAV 1022  
DB 641 WRHMFCTGTQHMGDPCVTALGASLCPSEQVAVSVICSGNQSLSSCNSSSLGPTPTI 700  
QY 1023 PEGSALICLEDKRLVDCDSRCAGRVEIYHDGFWGTCIDCDGWDLSDAHVVCQKLCGVA 1082  
DB 701 PEESAVACIESQQLRLVNGGRCAGRVEIYHGGSWGTCIDCDSDWDLSDAHVVCQKLCGGA 760  
QY 1083 FNATVSAHFEGSGPIWLDLNLCTGTESHLWQCPGSRGWGQHCRCRKEDAGVICSEFTALR 1142  
DB 761 INATGSAHFEGSGTPIWLDLNLCTGTESHLWQCPGSRGWGQHCRCRKEDAGVICSEFTALR 820  
QY 1143 LYSETETESACGRLEVFYNGTWGSGVRNRTTAAIGIVCQKLCGCGNGVVSAPLSKTS 1202  
DB 821 LTSEASREACAGRLEVFYNGTWGSGVRNRTTAAIGIVCQKLCGCGNGVVSAPLSKTS 880  
QY 1203 GPMWDDTDCPKTHISWCLSLAPWERRISSPAEETWITCEDRIVRVGGDTSCSGRVEIW 1262  
DB 881 IPMVDNVQCPKGPDTLWQCPSPWPWKRLASPEETWITCEDRIVRVGGDTSCSGRVEIW 940  
QY 1263 HAGSWGTCDDSDWDLAEAEVWCQQLGCGSALAALRDASFGQGTGTWLDMDRCCKNESFL 1322  
DB 941 HGGSWGTCDDSDWDLDAQVVCQQLGCGPALKAFKAEAFEGQGTGTWLDNEVKCKNESL 1000  
QY 1323 WCHAKPWGSCGCKHEDAGVRCQSGLK-----SLNASSGHLALILSSIFGLLLVLF 1376  
DB 1001 WDCPARRWGHSECGHEDAAVNCTDISVQKTPQKATGRSSQSSFIAGVILGVVLLAIF 1060  
QY 1377 I--LFTWCRVQOKHPLRVSTRRSGSLEENLFHEMETCKLKRED 1419  
DB 1061 VALFELTKRRQROR---LAVSSRGENLVHQIYQREMNNSCLNADD 1102

## RESULT 9

US-09-977-577-12  
; Sequence 12, Application US/09977577  
; Patent No. US2002015595A1  
; GENERAL INFORMATION:  
; APPLICANT: MOESTRUP, Soren  
; APPLICANT: MOLLER, Holger J.  
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR AND THE USE  
; FILE REFERENCE: MOESTRUP-1A  
; CURRENT APPLICATION NUMBER: US/09/977,577  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: US 60/270,120  
; PRIOR FILING DATE: 2001-02-22  
; PRIOR APPLICATION NUMBER: DK PA 2001 00039  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: DK PA 2000 01543  
; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 1156  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-977-577-12

Query Match 42.1%; Score 3429; DB 9; Length 1156;  
Best Local Similarity 56.2%; Pred. No. 7.1e-238;

Matches 598; Conservative 161; Mismatches 294; Indels 12; Gaps 4;

Qy	364	GADLELRADLADGNSNCSGRVEYRHEQWWTICDQNMKNKEQALVWCQKQJGCPSPVFGSSRAK	423
Db	41	GTDKEILRLVDGNNKCSGRVEYVQVEWGTVCNNGWMSAEVAVICQJGCPTAIKAPQGAN	100
Qy	424	PSNEARDIWNSTISCTGNSALWDCYDQKAKRT-CFRRSDAGYVCSDKADLDLURLVGAH	482
Db	101	SSAGSGRIWMDHVSCHRGNSALWDCKHGWGKHNSCTHQQDAGVTCSDGSLNLEWRLTRGG	160
Qy	483	SPCYGRLEVKYOGEWGTVDCHDRMSTNAAVVCKQJGCGKPMHVFMTYTFKEASGPIWLDD	542
Db	161	NMCSGRIEIKFOGRWGTVCDDNFNDHASVTCRQLECSGVSFSGSSNFGEGSGPIWFD	220
Qy	543	VSCIGNESNIWDEHSGWGHKNCVHREDVIVTCSGDATWGLRLVGGSNRCSGRLEVYFOG	602
Db	221	LICNGNESALNCKHGWGKHNDHAEADAGVTCRKAGDLSURLVDGVTCSGRLEVRFP	280
Qy	603	RWGTVCDDGNSKAAAVVCSQJDCPSSIIGMGLGNASTGYGKIWLDDVYDCGDDESJLWSC	662
Db	281	EWGTCDDGWDSYDAAVACKQLGCPATAVTAIGRVNASKGFHILWDSVSCOGHEPAYWOC	340
Qy	663	RNSGWNDCSHSEDEVGVTCSASDWELRLVGGSSRCAGKVEVWVQAGVILCANGCMN	722
Db	341	KHHEWKHYCNHNEADAGVTCSDGSOLERLRGGGSRCACTVEYIQRLLKGVKCDRGWGLK	400
Qy	723	IAEVCVCRQLECSGSAIRVSRPEHFTERTLHILMNSGCTGGASJLWDCIRNEMKOTACHLN	782
Db	401	EADVCRQJGCGSALKTSYQVYSKIQANTWJLFLSSCMNETSLWDCCKNMWOGJLTC	460
Qy	783	MEASLICSABRQRLVADGMPDCSGRVEYKHADTRSVDCDSFSLHAANVLCRELNCGD	842
Db	461	EBAKITCSABRRLVGGDIPCSGRVEYKHGDTWGSICDSFSLAASVLCRELQCGTVV	520
Qy	843	SLSVGBHFGKGNGLTWAERFQCEGSETHALALCPIVQHEDPTCIHSREVGVVCSRYTDVRL	902
Db	521	SIILGGAHFEGEGNQIWAEEFQCEGHESHLSCPVAPREPGTCSHSRDVGVVCSRYTIRL	580
Qy	903	VNGKSOCDOGVETNVLGHWSGLDTHWDPEDARVLCRLSCGTALSTYTGGRYIGERSVRV	962
Db	581	VNGKTPCEGRVELKTLGWSGLNSHWDIEDAHLVLCQJLKCQVALSTPGGARFKGNGOI	640
Qy	963	WGHRFHCLGNESLNNQMTVLGAPCTHGNVTYVICTGSLTQPLFPCLANWSDPYLSAV	1022
Db	641	WRHMFHCTEQRHMGDCPYTLGASLCPSQVASYVICSNGSQSTLSSCNSLGPTRPTI	700
Qy	1023	PEGSALICLEDKRLRLVDGSDRCAGRVEYHDGFWGTICDDGWDLSDAHVVCQRLGCGVA	1082
Db	701	PEESAVACIESQRLVNGGRCAGRVEYHGGSWGTLICDSDWDLSDAHVVCRLGCGCEA	760
Qy	1083	FNATVSAHFGEGGPTIWLDDLNCTGESHWQCPSRGNGQHDCHKEADAGVICSEFTALR	1142
Db	761	INATGSAHFEGEGTPTIWLDEMCKNGKESIRMOCHSHGWOQGNCRHKEDAGVICSEFMSLR	820
Qy	1143	LYSETETESACGRLEVYNGTWGSGRRNITTAIAGIVCRLGCGENGVWSLAPLSGTGS	1202
Db	821	LTSEASREACAGRLEVYNGANGTVGKSSMSETTVGVVVCRLGCADGKGINPASLDKMS	880
Qy	1203	GFMMVDDIOCPKTHISIWQCLSPABERRISSPAEETWTICEDRIRVRGGDTECSGRVEIW	1262
Db	881	IPMWVNVQCPKGPDTLWQCPSSPWEKRLASPEETWTICDNKIRLQEGPTSCSGRVEIW	940
Qy	1263	HAGSWGTVCDSDWDLAEAEVWCQJLGCSSALAAALRDAFSFGGTGTIWLDDMRCKGNESFL	1322
Db	941	HGSGSWGTVCDSDWDLDDAQVQCQJGCGPALKAPKEAEFGGTGTPIWLNEYKCKGNESL	1000
Qy	1323	WDCBAPWGOSDCGHKEADAGVCSGOSLK-----SLNASSGHIALILSSYFGLLLVLF	1376
Db	1001	WDCPARRGWHSECHKEADAANVCTDISVQKTPQKATTCGRSRQSFYAVGLGVVLLAIF	1060
Qy	1377	I--LFTWCRVQKQKHLPLRVSTRRRGSLSEENLFHEMETCLUKRED	1419
Db	1061	VALFFUTKKRRQROR---LAVSSRGENLVHQIQYREMNNSCLNADD	1102

```

RESULT 10
US-09-977-577-11
; Sequence 11, Application US/09977577
; Patent No. US2002015595A1
; GENERAL INFORMATION:
; APPLICANT: MOESTRUP, Soren
; APPLICANT: MOLLER, Holger J.
; TITLE OF INVENTION: THE FUNCTION OF A HAPTOGLOBIN-HAEMOGLOBIN RECEPTOR
; FILE REFERENCE: MOESTRUP-1A
; CURRENT APPLICATION NUMBER: US/09/977,577
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: US 60/270,120
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: DK PA 2001 00039
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: DK PA 2000 01543
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1149
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-977-577-11

```

Db 641 DIEDAHLVLCQKCGVALSTPGAREGKNGQIWRHMFHCTCTEHOHGDPCPTALGASLC 700  
QY 990 IHGNTVSVICTGSLTOPLPFCIANVSDPYLSAVPEGSALICLEDKRLRLVDGDSRAGRV 1049  
Db 701 PSEQVASVCSGNSQTLSSNSSSLGPTPTIPBESAVACTESQGLRLVNGGRCAGRV 760  
QY 1050 EYHDFGFWCTICDDGWDLSDAHVCQKLCGCVAFNATVSAHFEGSGPTWLDLACTGTE 1109  
Db 761 EYHESGWTICDDSDLSDAHVCQKLCGCVAFNATVSAHFEGSGPTWLDLACTGTE 820  
QY 1110 SHLWQPSRGWQHCHRHKEDAGVICTSEPTALRLYSETETESACRLEVFYNGTWSVGR 1169  
Db 821 SRIWOCHSHGWQONCRHREDAVIGCSSEPMRLTSEASREACAGRLEVFYNGWGTGK 880  
QY 1170 RNITTAIGIVCRQLCGGNGVVSAPLSTKSGPMWDDIOCPKTHISIWQCLSPWHER 1229  
Db 881 SSMSETVGVCRQLGCDKADKINPASLDKAMSPMWDVQCPKGPDTLWQCPSPWPK 940  
QY 1230 RISSPAETWITCEDIRVRGDTSCSGRVEIWHAGSWGTVCDDSDMLAEAEVVCQQLGC 1289  
Db 941 RLASPEETWITCDKIRLQEGTSCSGRVEIWHAGSWGTVCDDSDMLAEAEVVCQQLGC 1000  
QY 1290 GSALALRASQGTGTWLDLMDCKNGNESFLWDCHAKPWGSDCGHKEDEAGVRCGSGS 1349  
Db 1001 GPALKAFKEAFEGGTGPTWLNVEVKCKNGNESSLWDCPARRWGHSECGHKEDEAVNCTDIS 1060  
QY 1350 LK-----SLNASSGHLALISIFGLLLLVLF--LFTWCRVQKQKHLPLRVSTRRRG 1401  
Db 1061 VQTPKATGTRSSRSSFIAGVILGVLLAIFVALFLLTKRRQROR--LAVSSRGEN 1117  
QY 1402 SLEENLFHEMETCLKRED 1419  
Db 1118 LVHQIQYREMNNSCLNADD 1135

## RESULT 11

US-10-042-431-78  
; Sequence 78, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: FRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042,431  
; CURRENT FILING DATE: 2001-10-25  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 78  
; LENGTH: 1436  
; TYPE: PRT  
; ORGANISM: Bos sp.  
US-10-042-431-78

Query Match 37.9%; Score 3082; DB 9; Length 1436;  
Best Local Similarity 41.7%; Pred. No. 7.9e-213;  
Matches 613; Conservative 208; Mismatches 495; Indels 154; Gaps 23;

QY 28 CILLNSCFLLISFNGTDLERLVNGDPCSGRVEVVKFGQWGTVCDDGNNTASTIVCK 87  
Db 13 CVLLLG-----TWVGQALERLKDGVHRCGRVEVVKHQEGWTVDCGYRWTLDKASVVC 67  
QY 88 QLCGPFSFAM---FRFGQALVRGKTLWLDVSCYGNESALWEQCH---REWGSHNYCHGE 141  
Db 68 QLCGGAAGPFGGAYFGPGL---GPIWLLYTSCEGTSTVSDCEHNSIKDYRNDGYNHGR 124

QY 142 DGVNVCYGEANLGLRLVDGNNSCGRVEVVKFQERWGTICDDGNMNLNTAAVVCROLGCPSS 201  
Db 125 DAGVVCSSG----FVRLAGDGPSCGRVEVHSGEAMIPVSDGNFTLATQAIICAEIGCCGA 180  
QY 202 FISSGVNPAVLRPIWLDDILCOGNELALMNCRRHGNHDCSHNEDVTLTICYDSSDLE 261  
Db 181 VSVLGHFELRESSAQWAEERFCEGEPELWVCPVPGGTCHHSGSAQVVCVSAVSEVR 240  
QY 262 LELVGGTNRCAGRVRLKIOGRWGTVCHHKNNAADAVVCKOLGCGCTALHFAGLPHLQSSG 321  
Db 241 L-MTNGSSQCEGOVEMNISGWRALCACSHWSLANANAVICRQLGCGVATSTPGPHLVEEG 299  
QY 322 DVVWLDGVSCSNESFLMDCRHSGTVNFDCDQHNDVSVICSDGADLELRADGNNCSGR 381  
Db 300 DQILARFHCSGAESFLWNSCPVTALGGPDCSHGNTASVICSN----- 342  
QY 382 VEVRIHEQWTTICDQNWKNQALVVCQLGCPFSVFSRRKPSNEARDIWINISCTGN 441  
Db 343 -QIQVLPO-----CNDV-----SQTGSAAS----- 363  
QY 442 ESALWDCTYDGKAKRTCFRRSDAGVICSDKADLRLVGAHSPCYGRLEVKYQGEWGTVC 501  
Db 364 -----EDSAPYCSDSR--QLRLVDGGGCPAGRVEILDQSGWGTIC 401  
QY 502 HDRWSTRNAAVVCQLGCGKPMHVFGMTYFKKASGPILDDVSCICGNESNIHDCRSHGW 561  
Db 402 DGDWDLDDARVVCROLGCGEALNATGSAHFAGSGPIWLDNLNCTKESHVWRCPSRGW 461  
QY 562 KINCVRHREDIVITCSGDATWGLRLVGGSNCRSLRLEVFGQWRGTVCDDGNNSKAAAVVC 621  
Db 462 QHNCRRHKQDAGVICS--EFLALRWVSEDOOCAGWLEVFYNGTWSVCRNPMEDITVSTIC 519  
QY 622 SOLDCPSSITIGMLGNAST----GYGKIWLDVSDGDESDLSWCRNSGWNDCSHSED 677  
Db 520 RQLGCGDS----GTLNSSVALREGFRPQWVDRIQCRKTDTSILWQCPSPMNTNSCPKE 575  
QY 678 VGVICSDASDMELRLVGGSSRCAGKVEVNVQGAVGILCANGMGMNIAEYVVCRLCGSAI 737  
Db 576 ATYWCADS--RQIRLVDDGGRCGRVEILDQSGWGTICDDRWDLDDARVVCROLGCGEAL 633  
QY 738 RVSREPHTERTLHILMSNGCTGGEASLWDCIRWENKOTACHLNWEASLICASAHQRL 797  
Db 634 DATVSSFFGTGSGPIWLDVNCRGEESQVWRCPFWGVRQHNCHQEDAGVICSGF--VRL 691  
QY 798 VGADMEPCSRGVEVVKHADTWRSVCDSDFLHAANVLRELNCGDALSLVSDHFGKNGLT 857  
Db 692 AGDGPSCGRVEVHSGEAMTPVSDGNFTLPTAQVICAELCGCKKAVSLGHMFPRESDGQV 751  
QY 858 WAEKFOCEGSETHLALCPVQHPEDTCHSREYVGVVCSRYTDLRLV--NGKSCQDGOVEIN 916  
Db 752 WAEFRCDGGEPELWNSCPVPCPGGTCLHSGAAQVVCVSVTEVQLMKNCTSOCEGOVEMK 811  
QY 917 VLGHWSLCTDTHWDPEDARVLCRQLSCGTALSTGTGKYIGERSVVRWGHFRHCLGNESLL 976  
Db 812 ISGRWALCASHWSLANAVVCROLGCGVATSTPGPHLVEEGDQISTAQFHCSGAESFL 871  
QY 977 DNMCTVLGAPCIGHNTVSVICTGSLTOPLPFCIANVSDPYLSAVPEGSALICLEDKRL 1036  
Db 872 WSCPVTALGGPDCSHGNTASVICSNGHNTQVLPQCNDFLSQPAGSAASESSPYCSDSRQL 931  
QY 1037 RLVDGDSRCAGRVEIYHDFGFWGTICDDGWDLSDAHVVCKLGGCVAFNATVSAHFEGSG 1096  
Db 932 RLVDGCGPGCGRVEILDQSGWGTICDDWDLDDARVVCROLGCGEALNATGSAHFAGSG 991  
QY 1097 PTWLDLNTCTGTHLWQCPSRGWGQDCHRHKEDAGVICTSEPTALRLYSETETESACRGL 1156  
Db 992 PTWLDLNTCTGTHLWQCPSRGWGQDCHRHKEDAGVICTSEPTALRLYSETETESACRGL 1049  
QY 1157 EYFNGTWSVGRNNTTAAIGIVCRQLCGGNGVVSAPLSTKSGPMWDDIOCPKTH 1216  
Db 1050 EYFNGTWSVGRNNTTAAIGIVCRQLCGGNGVVSAPLSTKSGPMWDDIOCPKTH 1109  
QY 1217 ISIWQCLSPAPWERRISSPAETWITCE-----DRIRVRGDTCECSGRVE 1260

Db	1110	TSLWQCPSPGWKYS	CSPEEYI	SCGRRPKSCPTAACTDRKLRUGDSECSGRVE	1169
Qy	1261	IWHAGSWGTVCDSDWDLAEAEVVCQGLCGS	ALAAALRDASFQCGTGTI	ILWDDMRCKNES	1320
Db	1170	VHNGSWGTVCDSDWSLAEAEVVCQGLCGQ	ALAAVRAAFPGNGSI	TLWDEVQCGRES	1329
Qy	1321	FLWDCHAKPWGQDCGHKEDAGVRCSSG	-----GSLKSLNASSGSHLIL	1364	
Db	1230	SLWDCVAEPWGQSDCKHEEDAGVRCSGVRT	LTPTTAGTRTTSNLSPGISLPGVLCIL	1289	
Qy	1365	SSIFGGLLLVLFLFTWCRVQKQKHLPLURV	STRRRGSLENLPHENETCLKREDPHGTR	1424	
Db	1290	GSLLFLVLVLTVLTQLLRW	-RAER-----RALSSYEDALAEAYEELDYL	LTQKEGLGSP 1342	
Qy	1425	TS-DDTPNHGCEDASDTSLGLVPASEATK	1453		
Db	1343	DOMTDVPDENYDDAEVVPVPGTPSP	SGOGNE 1372		

RESULT 12

```

US-09-759-130B-448
: Sequence 448, Application US/09759130B
: Publication No. US2003002279A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: APPLICANT: McCarthy, Sean A
: APPLICANT: Fraser, Christopher C
: APPLICANT: Sharp, John D
: APPLICANT: Barnes, Thomas S
: APPLICANT: Kirt, Susan J
: APPLICANT: Mackay, Charles R
: APPLICANT: Myers, Paul S
: APPLICANT: Leiby, Kevin R
: APPLICANT: Wrighton, Nicolas
: APPLICANT: Goodearl, Andrew
: APPLICANT: Holtzman, Douglas A
: TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
: TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, AND OTHER
: TITLE OF INVENTION: USES.
: FILE REFERENCE: MPI00-5350MNM
: CURRENT APPLICATION NUMBER: US/09/759,130B
: CURRENT FILING DATE: 2002-09-16
: PRIOR APPLICATION NUMBER: US 09/479,249
: PRIOR FILING DATE: 2000-01-07
: PRIOR APPLICATION NUMBER: US 09/559,497
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 09/578,063
: PRIOR FILING DATE: 2000-05-24
: PRIOR APPLICATION NUMBER: US 09/333,159
: PRIOR FILING DATE: 1999-06-14
: PRIOR APPLICATION NUMBER: US 09/596,194
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 09/342,364
: PRIOR FILING DATE: 1999-06-29
: PRIOR APPLICATION NUMBER: US 09/608,452
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/393,996
: PRIOR FILING DATE: 1999-09-10
: PRIOR APPLICATION NUMBER: US 09/602,871
: PRIOR FILING DATE: 2000-06-23
: PRIOR APPLICATION NUMBER: US 09/420,707
: PRIOR FILING DATE: 1999-10-19
: NUMBER OF SEQ ID NOS: 460
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 448
: LENGTH: 1436
: TYPE: PRT
: ORGANISM: Bovine
US-09-759-130B-448

```

Query Match 37.9%; Score 3082; DB 9; Length 1436;  
Best Local Similarity 41.7%; Pred. No. 7.9e-213;

Matches	613;	Conservative	208;	Mismatches	495;	Indels	154;	Gaps	23;
Qy	28	CILLNSCFLISSFN	GTDLRLRLVNGDPCSGTVEVKFQGOWGTVCDDGNMTASTVVCK	87					
Db	13	CVLLLG----	TWYGGQALELRKDGVRHCRBGRVEYKHGEMGVTDGYRWTLKDSVVCR	67					
Qy	88	QLGPCFPFSAM---	PRFGVAOTRHGKIWDLVSCYCNSEALWECH---REWSHCNHYGE	141					
Db	68	QLGCCAATGPPGAY	FPGFL---GPWILLATSCGETESTVSDCEHSNIKDYNRDYNGNHR	124					
Qy	142	DVGNYNCYEANLGL	RILVDGNNSCSGRVEYKFOERWGTCICDDGNWLNTAAAVVCRQLGCPSS	201					
Db	125	DAGVYC-----	FVRLAGDGPDCSRGEVHSHGERAPIVPSDGNFTATAQTIAEALCGCKGA	180					
Qy	202	FISGGVNSPAVLRI	PWLDITLCOGNELALWNRHGWNHDCSHNEEDVTLTCTDYSSDLE	261					
Db	181	VSVLGHELFPRESS	AQVNAEEPCGEBEPBLWVCPVCPGGTCHHSSSAQVVCYSAYSEVR	240					
Qy	262	LRLVGGTRNRCMGR	VELKIQGRMGTVCHHKNNNAADVVYKQLGCCTLAHFAGLPHLOGS	321					
Db	241	L-MTNGSQCEGOV	EMNISGWRALCALCASHSLANANVICRLGCGVAISTPGGPHLVEEG	299					
Qy	322	DVWLDGVSCGNES	FLWDCRHSTVFNFCDLHQNDVSIVCSGDADLELRADGNNCSGR	381					
Db	300	QILITARFHCSAES	FLWSCPVTLALGGDPCHGHNTASVICSGN-----	342					
Qy	382	VEVRIHQWMTICD	NNKNEQALVYKQLGCPESVFGSRRAPKNEARDIWINSICTGN	441					
Db	343	-QIOLPQ-----	-----CNDV-----SQPTGSAAS-----	363					
Qy	442	ESALWDCITYDGK	KAKRTCFRRSDAGVICSADKLRLRLVGAHSPCYGRLEYKYOGEWGTYC	501					
Db	364	-----	-----EDSAPYCSDSR--QLRLVDGGGCPACAGRVEILDQGSWGTIC	401					
Qy	502	HDRSTRNAAVVCK	QCGGRPMHVFGMTYFKAESGPLWDDVSCIGNESNIWDCEHSGWG	561					
Db	402	DQDWLDLARVY	COLGCGEALNATGSAHFAGSGSPIWLNLNLTGKESHVMWRCSRGGW	461					
Qy	562	KHNCVHREDVTVC	SODATWGLRLVGGSNRCSGRLEYVFOGRMGTVCDDGNWNSAAAIVC	621					
Db	462	QHNCRHKKDAGV	ICS--EFFALRMVEDQQACAGWLEFYNGTWGVCVRNPMDITVSTIC	519					
Qy	622	SQDPCPSIIIGML	GNAST----GYGKIWLDDVSCDGEDSLDSCRSMGNNDCHSED	677					
Db	520	ROLCCGDS----	GLNSSVALREGFRPQWVDRIQCRKTDTSLWCPCSDPWNYNSCSPKEE	575					
Qy	678	VGVTCSADASDEL	RLVGGSSRCACKVEVNVOGAVGIILCANGWGNIAEVVCRQLECSCAI	737					
Db	576	AYIWCAADS--	RQIRLPDGGRCRGRVFDLQSGWGTICDDRWDLDARVYVCKQLGCGEAL	633					
Qy	738	FVSRPEHTERTLH	LSMNSCTCGEASLWDCIRWEWKOTACHLNMEASLICSAHRQPL	797					
Db	634	DATVSSFPGTGG	PIWDEVNCRGESSQVWHPCPSWGRQHMCNHQEDAGVICSGF--VRL	691					
Qy	798	VGADMPICSGRVE	KHADTWRSVCDSDPSFLHAANYLCRELNCGDIAISLVSGDFHKGNGLT	857					
Db	692	AGDGPCCSRVEH	SHGEANTVPVSDGNFTLPTAQVIAELGCGKAVSVLGHMPFPRESGQV	751					
Qy	858	WAEFQCEGESETH	ALCPVIQHPEDTCHSRREVGVCSRVDTRLV--NGKSQCQDQVQIN	916					
Db	752	WAEFRCDGGEPE	LWSCPRVPCGTCLHSGAAQVVCVSYTEVOLMKNGTSQCEGOVEMK	811					
Qy	917	VLGHWSICDTHPD	EADRVLCROLSCTALTSTGGXYIGERSVRVWVGRHFRHCLGNESLL	976					
Db	812	ISGRMARLCAASH	SLANANVVCROLGCGVAISTPRGPHLVEGGQIDTAQFHCSGAESFL	871					
Qy	977	DNCOMTVLAGAPP	CIHNTVSTIGTSLQTPLFPCLANVSDPYLSAVPEGGALICLEDKRL	1036					
Db	872	WSCPVTALGDC	SHGNTASVICSNHTQVLPQCNDFLSQPAGSAASEESPVCSDSRQL	931					
Qy	1037	RLVDGDSRCAGR	VEIYHDGFWGTICDDGDWLSDAHVVYCKLGGCGVAFNATVSAHFEGSG	1096					
Db	932	RLVDGGCGPCR	FEILDQGSWGTICDDWDLDARVYVCRQLGCGEALNATGSAHFAGAGSG	991					





Search completed: May 12, 2003, 06:47:08  
Job time : 162 secs

---

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 254 Seconds  
(without alignments)  
7108.859 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDGRCCCHQN.....CEDASDTSLLGLVPASEATK 1453

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cn2.1/USPTO.spool/US09759130/runat\_06052003.123555.23406/app\_query.fasta.1.1607  
-DB=Published\_Applications\_NA -QW=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09759130.@CGN.1.1.158.@runat\_06052003.123555.23406  
-NCPU=6 -ICPU=3 -NO\_XLPY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:

- 1: /cn2.6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cn2.6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cn2.6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cn2.6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cn2.6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cn2.6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cn2.6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cn2.6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cn2.6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cn2.6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cn2.6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cn2.6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cn2.6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cn2.6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	8138	100.0	4359	US-10-042-431-10 Sequence 10, Appl
2	8138	100.0	4359	US-09-759-130B-380 Sequence 380, Appl
3	8138	100.0	4628	US-10-042-431-9 Sequence 9, Appl
4	8138	100.0	4628	US-09-759-130B-379 Sequence 379, Appl

5	3082	37.9	4308	9	US-10-042-431-79	Sequence 79, Appl
6	3082	37.9	4308	9	US-09-759-130B-449	Sequence 449, Appl
7	1142.5	14.0	4344	10	US-09-917-800A-1712	Sequence 1712, Ap
8	830	10.2	2562	9	US-09-147-947-5	Sequence 5, Appl1
9	818	10.1	473	9	US-09-918-995-32438	Sequence 32438, A
10	722	8.9	2614	9	US-09-147-947-3	Sequence 3, Appl1
11	691.5	8.5	1686	9	US-09-905-291A-147	Sequence 147, App
12	691.5	8.5	1686	9	US-09-902-853-147	Sequence 147, App
13	691.5	8.5	1686	9	US-09-907-824-147	Sequence 147, App
14	691.5	8.5	1686	9	US-09-907-841-147	Sequence 147, App
15	691.5	8.5	1686	9	US-09-904-011-147	Sequence 147, App
16	691.5	8.5	1686	9	US-09-906-742-147	Sequence 147, App
17	691.5	8.5	1686	9	US-09-906-838-147	Sequence 147, App
18	691.5	8.5	1686	9	US-09-907-613-147	Sequence 147, App
19	691.5	8.5	1686	9	US-09-907-942-147	Sequence 147, App
20	691.5	8.5	1686	9	US-10-227-884-113	Sequence 113, App
21	691.5	8.5	1686	9	US-09-904-820-147	Sequence 147, App
22	691.5	8.5	1686	9	US-09-904-859-147	Sequence 147, App
23	691.5	8.5	1686	9	US-09-909-204-147	Sequence 147, App
24	691.5	8.5	1686	9	US-10-230-163-113	Sequence 113, App
25	691.5	8.5	1686	9	US-09-904-786-147	Sequence 147, App
26	691.5	8.5	1686	9	US-09-906-646-147	Sequence 147, App
27	691.5	8.5	1686	9	US-09-906-700-147	Sequence 147, App
28	691.5	8.5	1686	9	US-09-902-903-147	Sequence 147, App
29	691.5	8.5	1686	9	US-09-903-749A-147	Sequence 147, App
30	691.5	8.5	1686	9	US-09-903-786-147	Sequence 147, App
31	691.5	8.5	1686	9	US-10-218-631-113	Sequence 113, App
32	691.5	8.5	1686	9	US-10-230-338-113	Sequence 113, App
33	691.5	8.5	1686	9	US-09-902-736-147	Sequence 147, App
34	691.5	8.5	1686	9	US-09-904-119-147	Sequence 147, App
35	691.5	8.5	1686	9	US-09-904-956-147	Sequence 147, App
36	691.5	8.5	1686	9	US-09-907-794-147	Sequence 147, App
37	691.5	8.5	1686	9	US-10-230-414-113	Sequence 113, App
38	691.5	8.5	1686	9	US-09-902-692-147	Sequence 147, App
39	691.5	8.5	1686	9	US-09-903-520-147	Sequence 147, App
40	691.5	8.5	1686	9	US-09-904-462-147	Sequence 147, App
41	691.5	8.5	1686	9	US-09-905-056-147	Sequence 147, App
42	691.5	8.5	1686	9	US-09-907-925-147	Sequence 147, App
43	691.5	8.5	1686	9	US-09-904-553-147	Sequence 147, App
44	691.5	8.5	1686	9	US-09-905-381-147	Sequence 147, App
45	691.5	8.5	1686	9	US-09-905-381-147	Sequence 147, App

## ALIGNMENTS

RESULT 1  
US-10-042-431-10  
; Sequence 10, Application US/10042431  
; Publication No. US20020182675A1  
; GENERAL INFORMATION:  
; APPLICANT: MCCARTHY, Sean A  
; APPLICANT: BARNES, Thomas M  
; APPLICANT: FRASER, Christopher C  
; APPLICANT: SHARP, John D  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,  
; FILE REFERENCE: 10147-602  
; CURRENT APPLICATION NUMBER: US/10/042,431  
; PRIOR FILING DATE: 2001-10-25  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: Patent Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 4359  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-042-431-10

Alignment Scores: 0  
Pred. No.: 0  
Length: 4359

Score: 8138.00 Matches: 1453  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-759-130b-381 (1-1453) x US-10-042-431-10 (1-4359)

Qy 1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHisGlnAsn 20  
Db 1 ATGAGTGCCTCAAACTCGTGGCATATGATTTTGAAGATGCTGCTCATCAGAAC 60  
Qy 21 LeuPheSerAlaValThrCysIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuSer 40  
Db 61 CTTTCTCTGCTGCTAACTTGCATCTGCTCCCTCAATTCCTGCTCTCATCAGCAGT 120  
Qy 41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60  
Db 121 TTTAATGGAACAGATTGAGTTGAGGCTGGTCAATGAGAGCGGTCCCTGCTCTGGACA 180  
Qy 61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80  
Db 181 GTGAGGTGAATTCACGGACAGTGGGGACTGTGTGTGATGATGGTGAACACTACT 240  
Qy 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100  
Db 241 GCCTCAACTGCTGCTGCAACAGCTTGGATGCTCCATTTCTTTCGCCATGTTTCGTTTT 300  
Qy 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrGlyAsn 120  
Db 301 GGACAAAGCCGTGACATGACATGCAAAATTTGCTTGTATGATGTTCTCTTATGCAAT 360  
Qy 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGly 140  
Db 361 GAGTCAGCTCTCTGGGAATGTCACACCGGGAATGGGAAGCAATACTTTATCATGGA 420  
Qy 141 GluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160  
Db 421 GAAGATGTTGGTGAACCTGTTATGTTGGAAGCAATCTGGGTTTGGAGCTAGTGGATGA 480  
Qy 161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180  
Db 481 ACAACTCTGTTCAGGAGAGCTGGAGTGAATTTCCAAAGAAGGTGGGGACTATATGT 540  
Qy 181 AspAspGlyTrpAsnLeuAsnThrAlaValValCysArgGlnLeuGlyCysProSer 200  
Db 541 GATGATGGTGAACCTGAATCTGCTGCCGTGGTGTGCAGGCAACTAGGATGTCCATCT 600  
Qy 201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220  
Db 601 TCTTTTATTTCTCTGGAGTTGTTAATAGCCCTGTGTATTGCGCCCAATTTGGCTGGAT 660  
Qy 221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240  
Db 661 GACATTTATGTCAGGGGAATCAGTTGGCACTCTGGAATTCAGACATCGTGGATGGGA 720  
Qy 241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeu 260  
Db 721 AATCATGACTGCAGTCACAAATGAGGATGTACATTAACCTGTTATGATAGTAGTACTT 780  
Qy 261 GluLeuArgLeuValGlyClyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280  
Db 781 GAACAAAGCCTTGTAGTGGAACTAACCGCTGTATGGGAGAGTAGAGCTGAATAATCCA 840  
Qy 281 GlyArgTrpGlyThrValCysHisIleLysTrpAsnAsnAlaAlaAspValValCys 300  
Db 841 GCAAGTGGGGACCGTATGCCACCAATAGTGAACAATGCTGCACTGATGCTGATGCG 900  
Qy 301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320  
Db 901 AAGCAGTTGGGATGTGGAAACCGCACTTCACTTCCGTGGCTTCCTTCACTTTCAGTCAGG 960  
Qy 321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340

Db 961 TCTGATGTTGATGGCTTGATGGTGTCTCCGTCTCCGTGTAATGAATCTTTCTTTGGGAC 1020  
Qy 341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360  
Db 1021 TGCAGACATTCGGGAACCGTCAATTTTTCACCTGCTTCATCAAAACGATGTGTCTGTGATC 1080  
Qy 361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380  
Db 1081 TGCTCAGATGGAGCAGATTTGGAACCTGGAGCTAGCAGATGGAAGTAACAATTTGTTTCAGGG 1140  
Qy 381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400  
Db 1141 AGACTAGAGGTGAGAATTCATCAACAGTGGTGGACAATATGTGACCAACATGGAAGAAAT 1200  
Qy 401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420  
Db 1201 GAACAAGCCCTTGTGTTTGAAGCAGCTAGGATGCTCGTTCAGCTCTTTGGCAGTCGT 1260  
Qy 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440  
Db 1261 CGTGTAAACCTAGTAATGAAGCTAGACACATTTGGATAAAACAGCATATCTTGCACCTGG 1320  
Qy 441 AsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThrCysPheArg 460  
Db 1321 AATGAGTCAGCTCTCTGGAGCTGCACATATGATGAAAAGCAAGCAACATGCTTCCGA 1380  
Qy 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480  
Db 1381 AGATCAGATGCTGGAGTAATTTGTTCTGATGAAGCAGATCTGACCTAAGGCTTTCGGG 1440  
Qy 481 AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyCyluTrpGlyThrVal 500  
Db 1441 GCTCATAGCCCTGTATGGAGATTGGAGTGAATACCAAGGAGAGTGGGGAGCTGTG 1500  
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520  
Db 1501 TGTCTAGCAGATGAGGACCAAGGAATGCAGCTGTGTGTGTGTAACAATTTGGATGTGA 1560  
Qy 521 LysProMetHisValPheGlyMetThrThrPheLysGluAlaSerGlyProIleTrpLeu 540  
Db 1561 AAGCCTATGATGCTGTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGCTG 1620  
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560  
Db 1621 GATGACGTTTCTTGGATTGGAATCAGTCAAAATATCTGGAGCTGCAACACAGTGGATGG 1680  
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580  
Db 1681 GGAAGACATAATTTGTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGTGCATGCAACA 1740  
Qy 581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe 600  
Db 1741 TGGGCTTGGAGCTGGTGGCGGAGCAACCGCTGCTCGGAAGACTGGAGGTGACTTT 1800  
Qy 601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal 620  
Db 1801 CAAGACCGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCACGCTGGGTG 1860  
Qy 621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640  
Db 1861 TGTACCCAGCTGGACTGCCATCTTCTATCATTTGCGATGGGTCTGGAAACGCTTCTACA 1920  
Qy 641 GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660  
Db 1921 GGATATGAAAAAATTTGGCTCGATGATGTTTCTCTGTATGGAGATGAGTCAGATCTCTGG 1980  
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680  
Db 1981 TCATCAGGAACAGTGGGTGGGAAATTAATGACTGCAGTCACAGTGAAGATTTGGAGTG 2040  
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyLysSerSerArgCysAla 700

Db 2041 ATCTGTTCTGATCGATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100  
QY 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTTPGly 720  
Db 2101 GGAAGTTGAGGTGAATGTCAGAGGTGCGTGGGAATCTCTGTCTAATGCTGGGA 2160  
QY 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740  
Db 2161 ATGAACATTGCTGAAGTTGTTTCAGGCAACTTGAATGTGGGTCTGCAATCAGGGTCTCC 2220  
QY 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760  
Db 2221 AGAGGCTCATTTACAGAAAGAAATACACATCTTAATGTGGAATCTGGCTGCAC 2280  
QY 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780  
Db 2281 GGAGGGGAAGCCTCTCTCTGGATTGTATACGATGGAGTGGAAACAGAGCTGCTCAT 2340  
QY 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800  
Db 2341 TTAATATGAAGCAAGTTTGATCTGCTCAGCCACAGCCAGCCAGGCTGGTTGGAGCT 2400  
QY 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820  
Db 2401 GATATGCCCTGCTCTGGAGCTGTGAAGTGAACATGCACACATGGCGCTCTGTCTGT 2460  
QY 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840  
Db 2461 GATTCGATTTCTCTCTGATGTCGCAATGTCTGTGCAGAGAAATTAATTTGGAGAT 2520  
QY 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860  
Db 2521 GCCATATCTCTTCTGTGGAGATCACTTTGGAAAAGGAATGCTTAACCTTGGCGCGAA 2580  
QY 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880  
Db 2581 AGTTCAGGTGGAAGGAGTGAACATCACTTGCATTTGCCCATTTGTTCAACATCCG 2640  
QY 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspVal 900  
Db 2641 GAAGACACTTGTATCCACACAGAGAAAGTTGGAGTTGCTGTTCCCGATATACAGATGTC 2700  
QY 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyClnValGluIleAsnValLeuGlyHis 920  
Db 2701 CGACTTGTGAATGGCAATCCCAAGTGTGACGGCAAGTGGAGATCAACGTGCTTGGACAC 2760  
QY 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940  
Db 2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAGATGCCCGTGTCTATGCAGACAG 2820  
QY 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyLysIleGlyLysValArgSerVal 960  
Db 2821 CTCAGCTGTGGAGTGTCTCTCAACACAGAGGAAATATATGGAGAAAGAGTGT 2880  
QY 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980  
Db 2881 CGTGTGTGGGACACAGGTTTCAATGCTTAGGGAATCAGTCACTTCTGGATAACTGTCAA 2940  
QY 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000  
Db 2941 ATGACAGTCTTGGAGCACCCTCTCTATCCATGGAATACTGTCTGTGATCTGCACA 3000  
QY 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTrpLeuSer 1020  
Db 3001 GGAAGCTGACCCAGCCACCTTTTCCATGCTCGCAATGTATCTGACCCATATTTGCT 3060  
QY 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040  
Db 3061 GCAGTTCACAGGGCAGTGTCTTGTATGCTGTAGAGGACAAACGGCTCCCGCTAGTGGAT 3120  
QY 1041 GlyAspSerArgCysAlaGlyArgValGluIleThrHisAspGlyPheThrGlyThrIle 1060  
Db 3121 GGGGACAGCCGCTGTGCGGGGAGATAGATCTATCACGACGGCTTCTGGGCAACCATC 3180

QY 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080  
Db 3181 TGTGATGACGGCTGGGACCTGAGCATGCCACGTGTGTCTCAAAAGCTGGGTGTGGA 3240  
QY 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100  
Db 3241 GTGGCTTCAATGCCACGGTCTCTGTCTACTTTGGGGAGGGGTGAGGCCCATCTGGCTG 3300  
QY 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120  
Db 3301 GATGACCTGAACATGCACAGGAACGGAGTCCCACTTGTGGCAGTCCCTTCCCGGGCTGG 3360  
QY 1121 GlyClnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140  
Db 3361 GGCAGCACGACATGCAGGCACAAGGAGGACGAGGGGTCTCTGCTCAGAAATTCACAGCC 3420  
QY 1141 LeuArgLeuTrpSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTrp 1160  
Db 3421 TTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTCTGGGAGATTGGAAGTCTTCTAT 3480  
QY 1161 AsnGlyThrTrpGlySerValGlyArgAsnIleThrThrAlaIleAlaGlyIleVal 1180  
Db 3481 AACGGGACCTGGGGCAGCGTGGCAGGAGGAACATCACACAGCCATAGCAGGATTTGTG 3540  
QY 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200  
Db 3541 TGCAGGCAGCTGGCTGTGGGAGAAATGGAGTTGTACAGCTCGCCCTTTATCTAAGACA 3600  
QY 1201 GlySerGlyPheMetTrpValAspIleGlnCysProLysThrHisIleSerIleTrp 1220  
Db 3601 GGCTCTGTTTCATGTGGTGGATGACATTCAGTGTCTTAAACGCATATCTCCATATGG 3660  
QY 1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluThrTrpIle 1240  
Db 3661 CAGTGCCTGTCTGCCCATGGGACCGAAGAAATCTCCAGCCCGCAGGAGACCTGGATC 3720  
QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260  
Db 3721 ACATGTGAAGATAGAATAAGAGTGGTGGAGGAGACACCGAGTGTCTGGGAGAGTGGAG 3780  
QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280  
Db 3781 ATCTGGCACGACGCTCTCTGGGACAGTGTGTGATCACTCTGGGACCTGGCGGAGCGG 3840  
QY 1281 GluValValCysGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSer 1300  
Db 3841 GAAGTGTGTGTACGACGCTGGGCTGTGGCTCTCTGCTGCTGGCTGCGGTGAGGACGCTTCG 3900  
QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320  
Db 3901 TTTGGCCAGGGAACATGCAACCATCTGTTGGATGCATGCGGTGCAAGGAAATGAGTCA 3960  
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340  
Db 3961 TTTCTATGGGACTCTACGCCAAACCCCTGGGACAGAGTGTGCTGGGACACACAGAGAGAT 4020  
QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360  
Db 4021 GCTGGCTGAGGTGCTCTGGACAGTCTGAAATCACTGAAATGCTCTCTCAGGTCAATTTA 4080  
QY 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380  
Db 4081 GCATTTATTTATCCAGTATCTTTGGGCTCTCTCTCTGTTGTTTATTTCTATTCTC 4140  
QY 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArg 1400  
Db 4141 ACGTGTGCGGAGTTCAGAAAAACATCTGCCCTTCAGAGTTTCAACCCAGAGAGG 4200  
QY 1401 GlySerLeuGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420  
Db 4201 GGTTCCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCCTCAAGAGAGAGACCCA 4260



Db 1141 AGACTAGAGGTGAGAAATTCATGAACAGTGGTGGCAATAATATGTGACCAGAACTGGAAGAAAT 1200  
Qy 401 GluclnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420  
Db 1201 GAACAAGCCCTTGGGTTTGTAAAGCAGCTAGGATGTCGGTTCAGCGTCTTTGGCAGTCGT 1260  
Qy 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440  
Db 1261 CGTGCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACACCATATCTTGCACCTGGG 1320  
Qy 441 AsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThrCysPheArg 460  
Db 1321 AATGAGTCAGCTCTCTGGGAGTCGACATATGATGGAAGAAGCAAAAGCCAAATGCTTCCGA 1380  
Qy 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480  
Db 1381 AGATCAGATGCTGGAGTAATTTGTCGTAAAGCCAGATCTGGACCTTAAGGCTTGTGGG 1440  
Qy 481 AlaHisSerProCysTrpGlyArgLeuGluValLysTrpGlnGlyGluTrpGlyThrVal 500  
Db 1441 GCTCATAGCCCTCTTATGGGAGATTGGAGTGAATACCAAGAGAGTGGGGACTGTG 1500  
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly 520  
Db 1501 TGTCTACACAGATGGACACAAGAAATGCAGCTGTTGTTGTAACAATTTGGGATGTGA 1560  
Qy 521 LysProMetHisValPheGlyMetThrTyrrPheLysGluAlaSerGlyProIleTrpLeu 540  
Db 1561 AAGCCTATGATGTTTGGTATGACCTATTTTAAAGAAGCATCAGGACCTATTTGGGCTG 1620  
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560  
Db 1621 GATCAGCTTCTTGCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1680  
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580  
Db 1681 GGAAAGCATAATTTGTATACACAGAGAGGATGTGATTTAAACCTGCTCAGGTGATGCAACA 1740  
Qy 581 TrpGlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyrrPhe 600  
Db 1741 TGGGGCCTGAGGCTGGTGGGGCGGAGCAACCGCTGCTGGGAAGACTGGAGGTGATCTTT 1800  
Qy 601 GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal 620  
Db 1801 CAAGGACGGTGGGCACAGTGTGTGATGACGGCTGGACAGTAAGCTGCAGCTGTGGTG 1860  
Qy 621 CysSerGlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThr 640  
Db 1861 TGTAGCCAGCTGGACTGCCATCTTCTATCATTTGGCATGGGCTCTGGAAACGCTTCTACA 1920  
Qy 641 GlyTyrrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrp 660  
Db 1921 GGATATGGAAAAATTTGGCTCGATGATGTTTCCTGTATGGAGATGAGTCAGATCTCTGG 1980  
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680  
Db 1981 TCATGAGGAACACTGGGTGGGGAATAATGACTGCAGTCACTGAGATGTTGGAGTG 2040  
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700  
Db 2041 ATCTGTTCTGATGATCGGATATGGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGTGCT 2100  
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720  
Db 2101 GGAAAAAGTTGAGTGAATGTGCCAGGGTCCGCTGGGGAATTTCTGTGCTAATGGCTGGGGA 2160  
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740  
Db 2161 ATGACATGCTGAAGTTGTTTGCAGGCAACTTGAATGGGGTCTGCAATCAGGGTCTCC 2220  
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760

Db 2221 AGAGAGCCTCATTTTACAGAAAGAACATTTACACATCTTAATGTGCAATTTCTGGTGCAC 2280  
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780  
Db 2281 GGAGGGAAGCCTCTCTCTGGGATGTATACGATGGGAGTGGAAACAGACTGCGTGTCTAT 2340  
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800  
Db 2341 TTAATATGGAAGCAAGTTTGATCTGCTACGCCACAGCAGCCAGGCTGTTGGAGCT 2400  
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820  
Db 2401 GATATGCCCTGCTCTGGACGCTTTGAAGTGAACATGTCAGACACATGGCGCTCTGTCTGT 2460  
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840  
Db 2461 CATCTCATTTCTCTCTTCATGCTGCCANTGTGTCGTCAGAGAAATTAATTTGGGAGAT 2520  
Qy 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860  
Db 2521 GCCATATCTCTTTCTGGGAGATCACTTTGCAAAAGGAATGGTCTAACTTGGGCCCAA 2580  
Qy 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880  
Db 2581 AAGTTCCAGTGTGAAGGAGTGAACCTCACTTGCATTTATGCCCATTTGTTCAACATCCG 2640  
Qy 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrrThrAspVal 900  
Db 2641 GAACACACTGTTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCCCGATATACAGATGTC 2700  
Qy 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920  
Db 2701 CGACTTGTGAATGCAAAATCCAGTGTGACGGCAAGTGGAGATCAACGTCTTGGACAC 2760  
Qy 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940  
Db 2761 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAGAATGCCCGTGTCTATGACAGAC 2820  
Qy 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyLysTyrrIleGlyGluArgSerVal 960  
Db 2821 CTCAGCTGTGGAGCTGCTCTCTCAACCCAGAGAGAAATATATTTGGAAAGAAGTGT 2880  
Qy 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980  
Db 2881 CGTGTGGGGACACAGCTTTTCATTTAGGGAATGAGTCACITCTCGATAACTGTCAA 2940  
Qy 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000  
Db 2941 ATGACAGTTCTTGGAGCACCTCCCTGTATCCATGCCAATACTGTCTCTGTGATCTGCACA 3000  
Qy 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrrLeuSer 1020  
Db 3001 GGAAGCCTGACCCAGCCACTGTTTCCATGCCCTCGCAAAATGATCTGACCCATATTGTCT 3060  
Qy 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuLeuValAsp 1040  
Db 3061 GCAGTTCCAGAGGCGAGTGTCTTGTATCTGTCTAGAGGACAAACGGCTCCGCTAGTGAT 3120  
Qy 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrrHisAspGlyPheTrpGlyThrIle 1060  
Db 3121 GGGGACAGCCGCTGTGCGGAGAGTAGAGATCTATCAGCAGCGGCTTCTGGGGCACCATC 3180  
Qy 1061 CysAspAspGlyTyrrAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080  
Db 3181 TGTGATGACGGCTGGGACCTGAGCGATGCCACGCTGGTGTGTCAAAAGCTGGGCTGTGGA 3240  
Qy 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100  
Db 3241 GTGGCTTCAATGCGCACCGTCTCTGCTCACTTTGGGAGGGGTGAGGGCCCATCTGGGTG 3300  
Qy 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120  
Db 3301 GATGACCTGAACTGCACAGGAACGAGTGCACCTTGTGGCAGTGCCTTCCCGGGCTGG 3360



Db 628 TCTTTTATTTCTTCTGGAGTTGTTAAATAGCCCTGCTCTATTGCGCCCATTTGGCTGGAT 687  
Qy 221 AspileLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240  
Db 688 GACATTTTATGCCAGGGAATGAGTTGGCACTCTGGAAATTCGACACATCGTGGATGGGA 747  
Qy 241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCystYrAspSerSerAspLeu 260  
Db 748 AATCATGACTGCAGTCACAATGAGGATGTCACATTAACTTGTATGATAGTAGTATCT 807  
Qy 261 GluLeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280  
Db 808 GAACAAAGGCTGTAGGTGGAACCTAAACCGCTGTATGGGAGAGTAGAGCTGAAAAATCAA 867  
Qy 281 GlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCys 300  
Db 868 GGAAGGTGGGGGACCGTATGCCACCAATAGTGAACATGCTGAGCTGATGCTGATGC 927  
Qy 301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320  
Db 928 AAGCAGTTGGGATGTCAGCCGCACTTCCTTCGCTGGCTTCCTCATTTGCCACTCAGGG 987  
Qy 321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340  
Db 988 TCTCATCTGTATGGCTTGTATGGTGTCTCTGCTCCGCTAATGAATCTTTCTTTGGGAC 1047  
Qy 341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360  
Db 1048 TGCAGACATTCGGGAACCGTCAATTTTGACTGTCTTCATCAAAACGATGCTGTGTATC 1107  
Qy 361 CysSerAspGlyAlaAspLeuLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380  
Db 1108 TGCTCAGATGGAGCAGATTTGGAACCTGGACTACGACATGGAACATAAATTTGTCAGG 1167  
Qy 381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400  
Db 1168 AGACTAGAGGTGAGAAATTCATGAACAGTGGTGGCAATATGTGACCAGAACTGGAAGAA 1227  
Qy 401 GluGlnAlaLeuValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420  
Db 1228 GAACAACCCCTTGTGGTTGTAAACAGCTAGGATGTCGGTTCACCGCTCTTGGCAGTCGT 1287  
Qy 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440  
Db 1288 CGTCTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACACATATCTTGCACCTGG 1347  
Qy 441 AsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThrCysPheArg 460  
Db 1348 AATGAGTCAGCTCTCTGGGACTGCACATATGATGGAAAGCAAGCAACATGCTTCCGA 1407  
Qy 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480  
Db 1408 AGATCAGATGCTGGAGTAATTTGTTCATAGGACAGATCGGACCTTAAGGCTTGTCCGG 1467  
Qy 481 AlaHisSerProCysTrpGlyArgLeuGluValLysTrpGlnGlyGluTrpGlyThrVal 500  
Db 1468 GCTCATAGCCCTCTTATGGGAGATTGGAGGTGAATACCAAGCAGAGTGGGGGACTGTG 1527  
Qy 501 CysHisAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGly 520  
Db 1528 TGTTCATACAGATGGACACAAAGGAATGCAAGCTGTGTGTGTGAACAATTTGGGATGTTGA 1587  
Qy 521 LysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu 540  
Db 1588 AAGCCTATGCATGTGTTTGTATGACCTATTTTAAAGAGCATTCAGGACCTATTTTGGCTG 1647  
Qy 541 AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp 560  
Db 1648 GATCAGCTTCTTCCATTTGGAATGAGTCAAAATATCTGGGACTGTGAACACAGTGGATGG 1707  
Qy 561 GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr 580  
|||||

Db 1708 GGAAGCATAAATTTGTGTACACAGAGAGATGTGATTGTAACCTGCTCAGGTGATGCAACA 1767  
Qy 581 TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTrpPhe 600  
Db 1768 TGGGCGCTGAGGCTGTGGGGCGGAGCAACCGCTGCTCGGGAACAGTGGAGGTGACTTT 1827  
Qy 601 GlnGlyArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValVal 620  
Db 1828 CAAGCAGCGTGGGCGACAGTGTGTGATGACGCGCTGGAACAGTAAAGCTGCAGCTGTGTG 1887  
Qy 621 CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640  
Db 1888 TGTAGCAGCTGGGACTGCCCATCTTCATCATGTCATGCGCTGCGGAACCGCTTCTACA 1947  
Qy 641 GlyTrpGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrp 660  
Db 1948 GGATATGGAANAATTTGGCTCGATGATGTTTCTCTGATGAGAGATGACTCAGATCTCGG 2007  
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680  
Db 2008 TCATGACAGGAACACTGGGTGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067  
Qy 681 IleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAla 700  
Db 2068 ATCTGTTCTGATGCTCGATATGAGCTGAGGCTTGTGGGTGGAAGCAGGAGTGTGCT 2127  
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720  
Db 2128 GGAAGTTGAGGTGAATGTCAGGCTGCGGTGGGAATTCGTGTGCTTAATGGCTGGGA 2187  
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740  
Db 2188 ATGAACATTTGCTGAAGTTGTTTCAGGCAACTTGAATGCGGTCTGCAATCAGGCTCCTC 2247  
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760  
Db 2248 AGAGAGCTTCATTCACAGAAAGAACATACACATCTTAATGTCGAATTCGGCTGCACT 2307  
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780  
Db 2308 GGAGGGAAGCCTCTCTCTGGGATTTGATACGATGGGAGTGGAAACACAGCTCGGTGTCAT 2367  
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800  
Db 2368 TTAATAATGGAAGCAAGTTTGTATCTCAGCCACAGCCAGCCAGGCTGTTGGAGCT 2427  
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820  
Db 2428 GATATGCCCTCTCTCTGGACGTGTTGAAGTGAACATGACAGACACATGCGCTCTCTGT 2487  
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840  
Db 2488 GATTCTGATTTCTCTCTTCATGCTGCCAATGCTGTGTCACAGAAATTAATTTGGAGAT 2547  
Qy 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860  
Db 2548 GCCATATCTCTTCTCTGGGAGATCAGCTTGGAAAGGAATGCTTAACCTTGGGCCGAA 2607  
Qy 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880  
Db 2608 AAGTTTCAGTGTGAAGGGAGTGAACATCACCCTTGCAATATGCCCATTTGTTCAACATCCG 2667  
Qy 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspVal 900  
Db 2668 GAACACACTTGTATCCACAGCAGAGAACTTGGAGTTGTCTGTTCCCGATATACAGATGTC 2727  
Qy 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920  
Db 2728 CGACTTGTGAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAGAGTCAAGCTGTTGGACAC 2787  
Qy 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940  
Db 2788 TGGGGCTCACTGTGTGACACCCCACTGGGACCCAGAAAGATGCCCGTGTCTTATGACAGACAG 2847

Qy 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960  
Db 2848 CTCAGCTGGGAGTCTCTCAACACACAGGAGAAATATATGGAGAAAGTGT 2907  
Qy 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnGln 980  
Db 2908 CGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAAAGTCAA 2967  
Qy 981 MetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSerValIleCysThr 1000  
Db 2968 ATGACAGTCTTGGAGACCTCCCTGATCCATGCAATGGAATGATCTCTGTGATGACACA 3027  
Qy 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSer 1020  
Db 3028 GGAAGCTGACCCAGCCAGTGTTCATGCTCCATGCTCGCAATGATATCTGACCCATATTTGCT 3087  
Qy 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgValAsp 1040  
Db 3088 GCAGTTCCAGAGGCGAGTCTTGTATCTGCTTAGAGGACAAACGGCTCCGCTAGTGGAT 3147  
Qy 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060  
Db 3148 GGGACAGCCGCTGTGCGGGAGAGTACAGATCTATCAGGAGGCTTCTGGGGCACCATC 3207  
Qy 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080  
Db 3208 TGTGATGAGGCTGGGACCTGAGCGATGCCAGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267  
Qy 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100  
Db 3268 GTGGGCTTCAATGCCAGGCTCTCTCTCACTTGGGGAGGGGTGAGGCCCACTGGGTG 3327  
Qy 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120  
Db 3328 GATGACCTGAAGTGCACAGAACGAGTCCCACTTGTGGCAGTGCCTTCCCGCGCTGG 3387  
Qy 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140  
Db 3388 GGGACAGCAGCTGCAGGCACAAAGAGGAGCAGCGGGTCACTGCTCAGAAATTCAGACC 3447  
Qy 1141 LeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160  
Db 3448 TTGAGGCTCTACAGTGAAGTGAACAGAGAGTGTGTGTGGAGATGGAAGTCTTCAT 3507  
Qy 1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleVal 1180  
Db 3508 AACGGACCTGGGCGAGCTGCGCAGGAGGAACATCACACAGCCATAGCAGCATTTGT 3567  
Qy 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200  
Db 3568 TGCAGGCAGCTGGGCTGTGGGAGAAATGAGTTGTGAGCTCGCCCTTTATCTAAGACA 3627  
Qy 1201 GlySerGlyPheMetTrpValAspIleGlnCysProLysThrHisIleSerIleTrp 1220  
Db 3628 GCCTCTGGTTCTATGGTGTGATGACATTCAGTCTCTTAAACCATATCTCCATATGG 3687  
Qy 1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240  
Db 3688 CAGTGCCTGTCTGCCCTGAGGAGGAGAAATCTCCAGCCACAGAGAGACCTGGATC 3747  
Qy 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260  
Db 3748 ACATGTGAAGATAGAAATAGAGTGTGGAGAGACACCGAGTGTCTGGAGAGTGGAG 3807  
Qy 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280  
Db 3808 ATCTGGCAGCAGGCTCTCTGGGGACAGTGTGTGATGACTCTGGGACCTGGCCGAGGCG 3867  
Qy 1281 GluValValCysGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300  
Db 3868 GAAGTGGTGTGCACAGCTGGGCTGTGCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3927

Qy 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSer 1320  
Db 3928 TTTGCCAGGAACTGGAACTATCGTTGGATGATCGGTGCAAGGAAATGATCA 3987  
Qy 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340  
Db 3988 TTTCTATGGGACTGTACGCCAAACCCCTGGGACAGAGTGTGGACACAAAGGAAGAT 4047  
Qy 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360  
Db 4048 GCTGGCTGGAGTGTCTGGACAGTCTGTAATCACTCAATGCCTCCTCAGTCAATTA 4107  
Qy 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380  
Db 4108 GCACATATTTATTCAGTATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4167  
Qy 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArg 1400  
Db 4168 AGCTGGTCCGAGTTCAGAAACAAACATCTGCCCCCTCAGAGTTCACACAGAGGAGG 4227  
Qy 1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420  
Db 4228 GGTCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCTCAAGAGAGAGACCA 4287  
Qy 1421 HisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440  
Db 4288 CATGGACAAAGAACCTCAGATGACACCCCAACCTGTTGTGAAGATGCTAGCAGACACA 4347  
Qy 1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453  
Db 4348 TCGCTGTGGGAGTCT 4386

## RESULT 4

US-09-759-130b-379  
; Sequence 379, Application US/09759130B  
; Publication No. US2003002279A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: McCarthy, Sean A  
; APPLICANT: Fraser, Christopher C  
; APPLICANT: Sharp, John D  
; APPLICANT: Barnes, Thomas S  
; APPLICANT: Kist, Susan J  
; APPLICANT: Mackay, Charles R  
; APPLICANT: Myers, Paul S  
; APPLICANT: Leiby, Kevin R  
; APPLICANT: Wrighton, Nicolas  
; APPLICANT: Goodearl, Andrew  
; APPLICANT: Holtzman, Douglas A  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER  
; FILE REFERENCE: MPI00-5350MIM  
; CURRENT APPLICATION NUMBER: US/09759,130B  
; PRIOR FILING DATE: 2002-09-16  
; PRIOR APPLICATION NUMBER: US 09/479,249  
; PRIOR FILING DATE: 2000-01-07  
; PRIOR APPLICATION NUMBER: US 09/559,497  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 09/578,063  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: US 09/333,159  
; PRIOR FILING DATE: 1999-06-14  
; PRIOR APPLICATION NUMBER: US 09/596,194  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/342,364  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: US 09/608,452  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/393,996  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: US 09/602,871  
; PRIOR FILING DATE: 2000-06-23

Db	808	GAAC	TAA	GGCTT	AGGT	GGAACT	AAC	CCCTGT	AT	GGG	GAG	AGT	AGAA	TCAA	867							
Qy	281	Gly	Arg	Trp	Gly	Thr	Val	Cys	His	His	Tyr	Trp	Asn	Asn	Ala	Ala	Asp	Val	Val	Cys	300	
Db	868	GGAA	GGT	GGG	GAC	CGT	AT	GCC	ACC	ATA	AGT	TGA	GAC	AACT	GCT	GCA	CGT	GAT	GTC	GAT	927	
Qy	301	Lys	Gln	Leu	Gly	Cys	Gly	Thr	Ala	Leu	His	Phe	Ala	Gly	Leu	Pro	His	Leu	Gln	Ser	Gly	320
Db	928	AAG	CAG	TTG	GG	ATG	TGA	CCG	CAC	TTCA	CTT	CG	T	CG	T	GG	CTT	GC	CT	CA	TTG	987
Qy	321	Ser	Asp	Val	Val	Trp	Leu	Asp	Gly	Val	Ser	Cys	Ser	Gly	Asn	Gln	Ser	Phe	Leu	Trp	Asp	340
Db	988	TC	GA	TG	TTG	TAT	GG	CTT	GAT	GGT	GT	CCT	CT	CG	T	CG	T	TA	GA	AT	CT	1047
Qy	341	Cys	Arg	His	Ser	Gly	Thr	Val	Asn	Phe	Asp	Cys	Leu	His	Gln	Asn	Asp	Val	Ser	Val	Ile	360
Db	1048	TGC	AG	AT	CC	CGA	CCG	TC	AA	TTT	G	ACT	CT	CT	TC	AT	CA	AA	CG	AT	G	1107
Qy	361	Cys	Ser	Asp	Gly	Ala	Asp	Leu	Leu	Arg	Leu	Ala	Asp	Gly	Ser	Asn	Asn	Cys	Ser	Gly	380	
Db	1108	TG	CT	CA	AT	GG	AC	AT	TT	G	AA	CT	AG	CA	G	AT	GG	AA	GT	TA	CA	1167
Qy	381	Arg	Val	Glu	Val	Arg	Ile	His	Glu	Gln	Trp	Trp	Ile	Cys	Asp	Gln	Asn	Trp	Lys	Asn	400	
Db	1168	AG	AG	T	AG	AG	T	G	AG	AT	TC	AT	CA	AC	AG	T	GG	CA	AA	TA	T	1227
Qy	401	Glu	Gln	Ala	Leu	Val	Val	Cys	Lys	Gln	Leu	Gly	Cys	Pro	Phe	Ser	Val	Phe	Gly	Ser	Arg	420
Db	1228	GAC	AG	CC	CT	CT	GG	TTT	T	GT	T	AG	C	AG	C	T	AG	G	AT	GC	G	1287
Qy	421	Arg	Ala	Lys	Pro	Ser	Asn	Glu	Ala	Arg	Asp	Ile	Trp	Ile	Asn	Ser	Ile	Ser	Cys	Thr	Gly	440
Db	1288	CG	T	G	T	AA	CC	T	AG	T	AA	G	CT	AG	CA	AT	T	G	TA	AA	CC	1347
Qy	441	Asn	Gln	Ser	Ala	Leu	Trp	Asp	Cys	Thr	Tyr	Asp	Gly	Lys	Ala	Lys	Arg	Thr	Cys	Phe	Arg	460
Db	1348	AAT	G	AG	T	CA	GC	T	CT	GG	ACT	G	CA	CA	T	AT	G	AT	G	AA	AG	1407
Qy	461	Arg	Ser	Asp	Ala	Gly	Val	Ile	Cys	Ser	Asp	Lys	Ala	Asp	Leu	Asp	Arg	Leu	Val	Gly	480	
Db	1408	AG	AT	C	AG	AT	G	CT	CG	GA	T	AT	T	GT	CT	G	AT	AG	G	C	AG	1467
Qy	481	Ala	His	Ser	Pro	Cys	Tyr	Gly	Arg	Leu	Glu	Val	Lys	Tyr	Gln	Gly	Glu	Trp	P6ly	Thr	Val	500
Db	1468	G	CT	CA	T	AG	CC	CC	T	GT	T	AT	GG	AG	AT	GG	AG	T	GG	AG	AG	1527
Qy	501	Cys	His	Asp	Arg	Trp	Ser	Thr	Arg	Asn	Ala	Ala	Val	Cys	Lys	Gln	Leu	Gly	Cys	Gly	520	
Db	1528	TG	T	CA	T	CA	CA	AG	T	GG	AG	CA	AG	AA	T	CG	AG	T	GT	GT	GT	1587
Qy	521	Lys	Pro	Meth	His	Val	Phe	Gly	Met	Thr	Tyr	Phe	Lys	Glu	Ala	Ser	Gly	Pro	Ile	Trp	Leu	540
Db	1588	AG	CC	TA	T	GC	AT	G	T	GT	T	GT	AT	GA	AG	CA	AT	CA	AG	CA	CT	1647
Qy	541	Asp	Asp	Val	Ser	Cys	Ile	Gly	Asn	Glu	Ser	Asn	Ile	Trp	Asp	Cys	Glu	His	Ser	Gly	Trp	

Qy	321	SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp	340
Db	988	TCGTGATGTTGATAGGCTTGATGTTGCTCCTGCTCCGGTAATGAATCTTTCTTTGGGAC	1047
Qy	341	CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle	360
Db	1048	TGCAGACATTCGGGAACCGTCAAATTTTGACTGCTTCATCAAAACGATGTGCTGTGATC	1107
Qy	361	CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly	380
Db	1108	TGCTCAGATGGACAGATTGGAACATGCGACTAGCAGATGGAAGTAACAATTTGTCAGGG	1167
Qy	381	ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn	400
Db	1168	AGAGTAGAGGTGAGAAATTCATGAACAGTGGTGGCAAAATATGTGACCAAGCTGGAAGAAT	1227
Qy	401	GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValIlePheLysSerArg	420
Db	1228	GAACAAGCCCTTGTTGTTGTAAAGCAGCTAGAGATGTCGCTCAGCGCTCTTTGGCAGTCGT	1287
Qy	421	ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly	440
Db	1288	CGTGTAAACCTAGTAATGAAGCTAGAGACATTTGGATAAACAGCATATCTTGCACCTGGG	1347
Qy	441	AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg	460
Db	1348	AATGAGTACAGCTCTGGGACTGCACATATGATGAAAGCAAGCAACATGCTTCCGA	1407
Qy	461	ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuValGly	480
Db	1408	AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGGCTTGTCCGG	1467
Qy	481	AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrVal	500
Db	1468	GCATCATTAGCCCCCTGTTATGGAGATTGGAGGTGAATACCAAGAGAGATGGGGGACTGTG	1527
Qy	501	CysHisAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGly	520
Db	1528	TGTCATGACAGATGGACACAGGAATGCAGCTGTGTGTGTAACAANTGGAGTGTGA	1587
Qy	521	LysProMethHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu	540
Db	1588	AAGCCTATGATGTTGTTGGTATGACCTATTTTAAAGAGCATCAGGACCTATTTTGGCTG	1647
Qy	541	AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp	560
Db	1648	GATGACGTTTCTTGCAATGGAAATGAGTCCAAATATCTGGGACTGTGAACACATGGATGG	1707
Qy	561	GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr	580
Db	1708	GGAAAGCATAATTTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA	1767
Qy	581	TrpGlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyrPhe	600
Db	1768	TGGGGCCTGAGGCTGGTGGCGCGCAACACCCGCTGCTCGGGAAGACTGAGAGGTGTACTTT	1827
Qy	601	GlnGlyArgTrpGlyThrValCysAspAspGlyTTrpAsnSerLysAlaAlaValVal	620
Db	1828	CAAGGACGGTGGGGCACAGTGTGTGATGACGCTGGAACATGTAACGCTGCAGCTGTGGTG	1887
Qy	621	CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr	640

DQ	1048	JGAGACATITCCGGAACCGTCAAAATTGTGACTGTCTTCATCAAAACGAGTGCTGTGAGTGC
QY	361	CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly
DB	1108	TGCTCAGATGGACGACATTTGGAACCTGCCACTAGCAGATGGAAGTAACAATTTGTTCAAGG
QY	381	ArgValGluValArgIleHisGluGlnTrpThrIleCysAspGlnAsnTrpLysAsn
DB	1168	AGAGTAGAGGTGAGAATTATGAACACAGTGGTGGACAATAATGTACCAGAACTGGAAGAAT
QY	401	GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg
DB	1228	GAACAGCCCTTGTGGTTTTGTAAAGACGTAGGAATGTCGCTCAGCGCTTTGGCAGTGGT
QY	421	ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly
DB	1288	CGTGTCAAACCTAGTAATGAAGCTAGACACATTTGGATAAACAGCATATCTTGGACTGGG
QY	441	AsnGluSerAlaLeuTrpAspCysThrTyrrAspGlyLysAlaLysArgThrCysPheArg
DB	1348	PARGTGTACGCTCTGTGGACTGCACATATGATGGAANGAAGCAAGCAACATGCTTCCGA
QY	461	ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuValGly
DB	1408	AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTGGACCTAAGCGTTGTCCGG
QY	481	AlaHisSerProCysTyrrGlyArgLeuGluValLysTyrrGlnGlyGluTrpGlyThrVal
DB	1468	GCTCATAGCCCCCTGTATTGGAGATTGGAGGTGAATACCAAGAGAGATGGGGGACTGTG
QY	501	CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly
DB	1528	TGTCATGACAGATGGACACAAGGAATGCAGCTGTGTGTGTTAAACAANTGGAGTGTGA
QY	521	LysProMethisValPheGlyMetThrTyrrPheLysGluAlaSerGlyProIleTrpLeu
DB	1588	AAGCCTATGATGTGTGTGGTATGACCTATTTTAAAGAAAGCATCAGGACCTATTTTGGGTG
QY	541	AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp
DB	1648	GATGACGCTTCTTTCATTTGGAAATGAGTCCAATATCTGGGACTGTGAACACATGGATGG
QY	561	GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr
DB	1708	GGAAAGCATAAATTTGTACACAGAGAGGATGTGATTGTAACCTGCTCAGGTGATGCAACA
QY	581	TrpGlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyrrPhe
DB	1768	TGGGGCCTGAGGCTGGTGGCGCGCACACCCGCTGCTCGGGAAAGACTCGAGGTGTACTTT
QY	601	GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal
DB	1828	CAMGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTPAAAGCTGCAGCTGTGGTG
QY	621	CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr

QY	361	CysSerAspArgLysAspLeuGlnLeuValGlnTyrTrpThrIleCysAspGlnAsnTrpLysAsn	400
Db	1108	TGCTCAGATGGACAGATTGGAACTGCGACTAGCAGATGGAAAGTAACAAATTTGTTTCAGGG	1167
QY	381	ArgValGluValArgIleHisGluGlnTyrTrpThrIleCysAspGlnAsnTrpLysAsn	400
Db	1168	AGACTAGAGGTGAGAAATTCATGAACAGTGGTGGACAATATGTGCACCAACTGGAAGAAT	1227
QY	401	GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValIlePheGlySerArg	420
Db	1228	GAAACAGCCCTTGTTGTTGTAAAGCAGCTPAGGATGTCGCTTCAGCGCTCTTTGGCAGTCGT	1287
QY	421	ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly	440
Db	1288	CGTGCTAAACCTAGTAATGAAGCTAGACATTTGGATAAAACAGCATATCTTGCACCTGG	1347
QY	441	AsnGluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArg	460
Db	1348	AATGAGTCAGCTCTCTGGGACTGCACATATGATGAAAAAGCAAGCAACATGCTTCGGA	1407
QY	461	ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly	480
Db	1408	AGATCAGATGCTGGAGTAATTTGTTCTGATAAGGCAGATCTCGACCTAAGCGTTGTCCGG	1467
QY	481	AlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyLysTrpGlyThrVal	500
Db	1468	GCTCATAGCCCCCTTTATGGGAGATTGGAGGTGAATACCAAGAGAGATGGGGGACTGTG	1527
QY	501	CysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGly	520
Db	1528	TGTCATGACAGATGGACACAAAGAAATGCAGCTGTGTGTGTAACAAATTCGGATGTGGA	1587
QY	521	LysProMethHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeu	540
Db	1588	AAGCCTATGCATGTGTTGGTATGACCTATTTTAAAGAAAGCATCAGGACCTATTGGCTG	1647
QY	541	AspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrp	560
Db	1648	GATGACGCTTTCTTGCAATGGAAATGAGTCCAAATATCTGGGACTGTGAACACAGTGGATGG	1707
QY	561	GlyLysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThr	580
Db	1708	GGAAAGCAATAATTGTGTACACAGAGAGGATGTGATTGAACCTGCTCAGGTTGATGCAACA	1767
QY	581	TrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPhe	600
Db	1768	TGGGGCCTGAGGCTGGTGGGGCGGCACCAACCGCTGCTCGGGGAAGACTGGAGGTGTACTTT	1827
QY	601	GlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValVal	620
Db	1828	CAAGGACGGTGGGGCACAGTGTGTGATGACGGCTGGAACAGTAAAGCTGCAGCTGTGGTG	1887
QY	621	CysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr	640

Db 1888 TGTAGCCAGCTGGAGCTGCCCATCTTCTATCATTTGGCATGGGTCTGGGAAACGGCTTCTACA 1947  
Qy 641 GlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrp 660  
Db 1948 GGATATGGAAAATTTGGCTCGATCATGTTTCTCTGTGATGGAGATGAGTCAGATCTCTGG 2007  
Qy 661 SerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyVal 680  
Db 2008 TCATGCAGGAACAGTGGGTGGGGAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTG 2067  
Qy 681 IleCysSerAspAlaSerAspMetCysLeuArgLeuValGlyGlySerSerArgCysAla 700  
Db 2068 ATCTGTTCTGATGATCGATGATGAGCTGAGGCTGTGGGTGGGAGCAGGAGTGTGCT 2127  
Qy 701 GlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGly 720  
Db 2128 GGAAAAGTTGAGTGAAATGTCACGGGTCCCTGGGAATCTGTGTCTGTAATGGCTGGGA 2187  
Qy 721 MetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer 740  
Db 2188 ATGAACATTGCTGAAGTTGTTGGCAGGCACTTGAATGTGGGTCTGCAATCAGGCTCTCC 2247  
Qy 741 ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThr 760  
Db 2248 AGAGAGCCCTCATTTTCACAGAAAGAACATTACACATCTTAATCTCGAATCTCGGCTGCACT 2307  
Qy 761 GlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHis 780  
Db 2308 GGAGGGGAAGCCTCTCTCTGGATTGTATACGATGGGAGTGGAAACACACTGCGTGTCTAT 2367  
Qy 781 LeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAla 800  
Db 2368 TTAATAATGGAAGCAAGTTGATCTGCTCAGCCACAGCCAGCCAGGCTGTTGGAGCT 2427  
Qy 801 AspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCys 820  
Db 2428 GATATGCCCTCTGCTGGAGCTGTTGAAAGTGAACATGACAGACATGCGCTCTGCTGT 2487  
Qy 821 AspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAsp 840  
Db 2488 GATTCGATTCTCTCTCATGCTGCCAATGTGCTGTGCAGAGAAATTAATTTGTGGAGAT 2547  
Qy 841 AlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGlu 860  
Db 2548 GCCATATCTCTTCTGCTGGAGATCATTGGAAAAGGAATGGTCAACTTGGGCCCAA 2607  
Qy 861 LysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisPro 880  
Db 2608 AAGTTCCAGTGTGAAGGAGTGAACCTCACCTTGATATGCCCATTTGTTCAACATCG 2667  
Qy 881 GluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTrpThrAspVal 900  
Db 2668 GAAGACACTGTATCCACAGCAGAGAAGTTGGAGTTGCTGTTCCTCCGATATACAGATGTC 2727  
Qy 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920  
Db 2728 CGACTGTGAATGGGAATCCCACTGTCACGGGCAAGTGGAGATCAACGTGCTGGACAC 2787  
Qy 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGln 940  
Db 2788 TGGGGCTCACTGTGTGACACCCACCTGGGACCAGGAAGATGCCCGTGTCTATTCAGACAG 2847  
Qy 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerVal 960  
Db 2848 CTCAGCTGTGGGACTGCTCTCTCAACCACAGGAGGAAAATAATATTGGAGAAAAGTGT 2907  
Qy 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980  
Db 2908 CGTGTGTGGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTTCTGGATAACTGTCAA 2967  
Qy 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000  
Db 2968 ATGACAGTTCTTGGAGCAGCTCCCTGTATCCATGGAAATACTGTCTCTGTGATCGACA 3027

Qy 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrIleuSer 1020  
Db 3028 GGAAGCCTGACCCAGCCACTGTTTCCATGCTCGCAATGTATCTGACCCCATATTTGTCT 3087  
Qy 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAsp 1040  
Db 3088 GCAGTCCAGAGGCGAGTGTCTTGTATCTGCTTAGAGACAAACGGCTCCGCTAGTGGAT 3147  
Qy 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIle 1060  
Db 3148 GGGACAGCCGCTGTCCGGGAGAGTAGAGATCTATCACGACGCTTCTGGGGCACCATC 3207  
Qy 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080  
Db 3208 TGTGATACGGCTGGGACCTGAGCGATGCCACGTGTGTCTCAAAAGCTGGGCTGGGA 3267  
Qy 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeu 1100  
Db 3268 GTGGCCTTCAATGCCCGCTCTCTGCTACCTTTGGGAGGGGTGAGGGCCCATCTGGCTG 3327  
Qy 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120  
Db 3328 GATGACCTGAACCTGCACAGGAACGAGTCCACCTTGTGGCAGTCCCTTCCCGCGCTGG 3387  
Qy 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140  
Db 3388 GGGCAGCAGCACTGCAGGCACAGGAGCAGCGGGTCTATCTGCTCAGAAATTCACAGCC 3447  
Qy 1141 LeuArgLeuTyrSerGluThrGluThrCysAlaGlyArgLeuGluValPheTyr 1160  
Db 3448 TTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTGCTGGGAGATGGAAGTCTTCTAT 3507  
Qy 1161 AsnGlyThrTrpGlySerValGlyArgAsnIleThrAlaIleAlaGlyIleVal 1180  
Db 3508 AACGGACCTGGGCGAGCTGCGCAGGAGGAACATCACACAGCATACGAGGCAATTGTG 3567  
Qy 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200  
Db 3568 TGCAGGCACTGGGCTGTGGGAGAAATGGAGTTGTACGCCCTCGCCCTTTATCTAAGACA 3627  
Qy 1201 GlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrp 1220  
Db 3628 GGCTCTGTTTCACTGTGGGTGATGACATTCACTGTCTCTAAACGCATATCTCCATATGG 3687  
Qy 1221 GlnCysLeuSerAlaProTrpGluArgIleSerSerProAlaGluGluThrTrpIle 1240  
Db 3688 CAGTGCCTGTCTGCCCATGGGAGGAAAGAAATCTCCAGCCAGCAAGAGACCTGGATC 3747  
Qy 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260  
Db 3748 ACATGTGAAGATAGAATAAGAGTGGGTGGAGAGACACCGAGTGTCTGGGAGAGTGGAG 3807  
Qy 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280  
Db 3808 ATCTGGCAGCAGGCTCTCTGGGGACAGTGTGTGATGACTCTCTGGGACCTGGCCGAGGG 3867  
Qy 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSer 1300  
Db 3868 GAACTGGTGTGTACAGCAGCTGGGCTGTGGCTCTGCTGTGGCTGCCCTGAGGAGCCTTCG 3927  
Qy 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320  
Db 3928 TTTGGCCAGGGAACCTGGAAACCATCTGGTTGATGATACATGCGGTGCGAAAGGAATGAGTCA 3987  
Qy 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340  
Db 3988 TTTCTATGGGACTGTACGCCAAACCTGGGACAGAGTGTGTCGACACAAGAGAT 4047  
Qy 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360  
Db 4048 GCTGGCGTGGAGTGTCTGGAGTGGAGTGGTGAATCACTGAATGCCTCTCCTCAGTCAATTTA 4107

262	Db	262	-----GGCCCCATTTGGCTTTTGTATATCTATCTGTAAGGACAGAGCTCAACTGTC	312	...
125	Qy	125	TrpGluCysGlnHis-----ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu	141	...
313	Db	313	AGTGACTGTGAGCATCTCTAATAATTAAGACTATCTGTAATGATGGCTATAATCATGGTCGG	372	...
142	Qy	142	AspValClyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspClyAsn	161	...
373	Db	373	GATGCTGGAGTAGTCTGCTCAGGA-----TTTGTGCGTCTGGCTGGAGGGAT	420	...
162	Qy	162	AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp	181	...
421	Db	421	GGACCTGTCTCAGGGCAGTAGAAGTGCATTTCTGGAGAAGCTTGGATCCCACTGCTCAT	480	...
182	Qy	182	AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer	201	...
481	Db	481	EGGAACCTTACACTGCCACTGCCAGATCATCTGTGCAGAGTGGGTGTGGCAAGCT	540	...
202	Qy	202	PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp	221	...
541	Db	541	GTGTCGTCTGGACATGAGCTCTTCACAGAGTCCAGTCCCGCAGGCTCGGCTCAAGAG	600	...
222	Qy	222	IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn	241	...
601	Db	601	TTCAGGTGTGAGGGGAGGAGCGCTGAGCTCTGGGTCTGCCCCAGAGTCCCTGTGCCAGG	660	...
242	Qy	242	HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu	261	...
661	Db	661	GGCAGGTGTCACCACAGTGGATCTGCTCAGGTGTGTTGTTTCAGCATCTCAGAAGTCCG	720	...
262	Qy	262	LeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly	281	...
721	Db	721	CTC---ATGACAAACGGCTCCTCTCAGTCTGAAGGCGAGTGCAGATGAACATTTCTGGA	777	...
282	Qy	282	ArgTrpGlyThrValCysHisIstLysTrpAsnAsnAlaAlaAspValValCysLys	301	...
778	Db	778	CAATGGAGAGCGTCTGTGCTCCCTCCACTCGAGTCTGGCCAAATGCCAAATGTTATCTG	837	...
302	Qy	302	GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer	321	...
838	Db	838	CAGCTCGGCTGTGGAGTTGGCATCTCCACCCCGGAGGACACACTTGGTGGAGAAGGT	897	...
322	Qy	322	AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys	341	...
898	Db	898	GATCAGATCTTAACAGCCCGATTTCATCTCTCTGGGCTGAGTCTCTCTGTGGAGTTGT	957	...
342	Qy	342	ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys	361	...
958	Db	958	CCTGTGACTCCCTGGGTGGTCTGCTACTGTTCCCATGGCAACACACGCTCTGTATCTGC	1017	...
362	Qy	362	SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg	381	...
1018	Db	1018	TCAGGAAC-----	1026	...
382	Qy	382	ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu	401	...
1027	Db	1027	--CAGATCCAGGTGCTTCCCCAG-----	1047	...
402	Qy	402	GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg	421	...
1048	Db	1048	-----TGCAACAGCTCCGTG-----	1062	...
422	Qy	422	AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrClyAsn	441	...
1063	Db	1063	TCCAACCTACAGGCTCGCGCTCA-----	1089	...
442	Qy	442	GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg	461	...
1089	Db	1089	-----	1089	...
462	Qy	462	SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla	481	...

Db 1090 GAGGACAGCGCCCTACTGCTCAGACAGCAGG-----CAGCTCCGCGCTGGTGAGCGG 1143  
Qy 482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501  
Db 1144 GCGGCTCCTCGCGGGAGAGTGGAGATCCTTGACAGGGCTCTCGGGGACCATCTGT 1203  
Qy 502 HisAspArgTrpSerThrArgAsnAlaValValCysLysGlnLeuGlyCysGlyLys 521  
Db 1204 GATGACGCGCTGGACGCTGACGATGCCGCGTGTGTGTCAGGACAGCTGGGTGGAGAA 1263  
Qy 522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541  
Db 1264 GCCCTCAATGCCAGGGCTGCTCCTCCTCGGCGCAGGATCAGGCGCCCATCTGTTGAC 1323  
Qy 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561  
Db 1324 AACCTGAACGCACAGAAAGAGTCCACAGTGTGGAGGTGCCCTCCCGGGCTGGGG 1383  
Qy 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581  
Db 1384 CAGCACAACTGCAGACACAAAGCAGCGGGGTCTCTGCTCA-----GAGTTCTCTG 1437  
Qy 582 GlyLeuArgLeuValGlyCysSerAsnArgCysSerGlyArgLeuGluValTyrPheGln 601  
Db 1438 GCCCTCAGATGGTGTGAGTGGAGACAGCAGTGTGTGGTGTGGAGTGTTCCTACAA 1497  
Qy 602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCys 621  
Db 1498 GGGACCTGGGCGAGTGTGCGTAACCCATGGAAGACATCACTGTGTCCAGCATCTGC 1557  
Qy 622 SerGlnLeuAspCysProSerSerIleleGlyMetGlyLeuGlyAsnAlaSerThr--- 640  
Db 1558 AGACAGCTTGGTGTGGGGACAGT-----GGAACCCCTCAACTCTCTTGTCT 1605  
Qy 641 -----GlyTyrGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSer 657  
Db 1606 CTTAGAGAGGTTTATAGCCACAGTGGTGGATAGATACCACTCATGCTTCCAAAGGAGAA 1665  
Qy 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677  
Db 1666 TCTCTCTGGCAGTCTTCTGACCTTGAATTAACAACATCATGCTTCCAAAGGAGAA 1725  
Qy 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyCysSerSer 697  
Db 1726 GCCTATATCTGGTGTGCAGACAGC-----AGACAGATCCGCTGGTGGATGGAGGTGGT 1779  
Qy 698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717  
Db 1780 CGCTGCTCTGGAGAGTGGAGATCCTTGACAGGGCTCTCGGGGACCATCTGTGATGAC 1839  
Qy 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGlyCysGlySerAlaIle 737  
Db 1840 CGCTGGGACCTGGAGATGCCGCTGTGTGTGCAAGCAGCTGGGCTGTGGAGAGCCCTG 1899  
Qy 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757  
Db 1900 GAGCGCACTGCTCTCTCTCTCTGGGACGGGATCAGGCGCCCATCTGCGTGAAGTG 1959  
Qy 758 GlyCysThrGlyGlyAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777  
Db 1960 AACTGCAGAGAGAGAGTCCCAAGTATGAGGTGCCCTCTCTGGGATGGCGGCAACAC 2019  
Qy 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797  
Db 2020 AACTGCAATCATCAAGAAGATCAGGAGTCACTGCTCAGGATTT-----GTGGCTCTG 2073  
Qy 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817  
Db 2074 GCTGGAGAGATGGACCTCTGCTCAGGGCGAGTAGAAGTGCATCTCGAGAGCCCTGGAC 2133  
Qy 818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837  
Db 2134 CCAGTGTCTGTGGAACCTTCACATCTCCCATCTGCCAGGTCTATCTGTGCAGAGCTGGGA 2193

Qy 838 CysGlyAspAlaIleSerLeuSerValCysAspHisPheGlyLysGlyAsnGlyLeuThr 857  
Db 2194 TGTGGCAAGGCTGTGTCTCTGGGACACATGCCATTCAGAGATCGGATGGCCAGGTC 2253  
Qy 858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877  
Db 2254 TGGGCTCAAGAGTTTCAGGTGTGATGGGGGAGGCTGAGCTGTCTGCTGCCCCAGAGTG 2313  
Qy 878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgTyr 897  
Db 2314 CCCTGTCCAGAGGACACATGCTCCACAGTGGAGCTGTCTGAGTGTCTGTCTGCTGATC 2373  
Qy 898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValIleLeuAsn 916  
Db 2374 ACAGAGTCCAGCTTATGAAAAACGGCACCTCTCAATGTGAGGGCAGGTGGAGATGAAG 2433  
Qy 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936  
Db 2434 ATCTGTGACGATGGAGAGCGCTCTGTGCCCTCCACATGGAGTGTGGCCCAATGCTT 2493  
Qy 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly 956  
Db 2494 GTCGTCTGACGCTGGCTGTGGAGTGGCATCTCCACCCAGAGGACACACTTGGTG 2553  
Qy 957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976  
Db 2554 GAAGAGGTGATCAGATCTCAACAGCCCAATTCACCTGCTCAGGGGCTGAGTCTCTCTG 2613  
Qy 977 AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996  
Db 2614 TGGAGTTCTCTGTGACTGCTTGGTGGGCTGACTGTCCCATGGCAACACACCTCT 2673  
Qy 997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016  
Db 2674 GTGATCTCTCAGAAACACACACCCAGGTGCTGCCCCAGTGCACAGACTTCTCTGCTCAA 2733  
Qy 1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036  
Db 2734 CCTGAGGCTCTCGGCGCTCAGAGAGAGTCTCCCTACTGCTCAGACAGCAGGAGGCTC 2793  
Qy 1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPhe 1056  
Db 2794 CCCTGTGTGGAGGGGGCGGCTCCCTCGCGGGGAGAGTGGAGATCTTTCACAGGCTCC 2853  
Qy 1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076  
Db 2854 TGGGCACTCTGTGATGATGACTGGGACCTGGAGATGCCCGTGTGGTGTGCAGGCGAG 2913  
Qy 1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096  
Db 2914 CTGGGCTGTGGAGAGCCCTCAATGCCACGGGTCTGCTCACTTCGGGCGAGATCAGGG 2973  
Qy 1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116  
Db 2974 CCCTCTGGCTGGAGCAGCTGACCTGACAGGAAAGAGTCCACGCTGTGGAGGTGCCCT 3033  
Qy 1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136  
Db 3034 TCCCGGGCTGGGGCGGCAGCTGACAGACACAAAGAGGAGCGCGGGTCTATCTGCTCA 3093  
Qy 1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156  
Db 3094 GAGTTCTCTGGCCCTCAGGATGGTGAGC-----GAGGACACAGAGTGTCTGGGTGGCTG 3147  
Qy 1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIle 1176  
Db 3148 GAGGTTTCTACACAGCGGACCTGGGGCAGTGTCTCCCGCAGGCCCCCATGGAATATCACT 3207  
Qy 1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196  
Db 3208 GTGTCCGTGATCTGCAGACAGCTTGGATGTGGGGACAGTGGAGTCTCAACACCTCTGTT 3267



Db 373 GATGCTGGAGTAGTCTGCTCAGGA-----TTTGTGCGGTCTGGCTGGAGGGAT 420  
Qy 162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181  
Db 421 GGACCTCTGCTCAGCGGAGTAGAGTGCATCTTGGAGAAGCTTGGATCCAGTGTCTGAT 480  
Qy 182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201  
Db 481 GGGAACTTTCACACTTGCCTGCCAGATCATCTTGCAGAGATTGGTGTGGCAAGGCT 540  
Qy 202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221  
Db 541 GTGCTGTCTCTGGGACATGAGCTCTTCAGAGAGTCCAGTGGCCAGGTCTGGGCTGAAGAG 600  
Qy 222 IleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241  
Db 601 TTCAGGTGTGAGGGGAGGAGCTGAGCTGCTGCCCTGCCAGAGTCCCTGTCCAGG 660  
Qy 242 HisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGlu 261  
Db 661 GGCAGTGTCCACAGAGTGTCTGCTCAGGTTGTTTTCAGCATATCTCAGAGTCCGG 720  
Qy 262 LeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281  
Db 721 CTC----ATGACAAACGGCTCTCTCAGTGTGAAGGCGAGGTGGAGATGAACATTTCTGGA 777  
Qy 282 ArgTrpGlyThrValCysHisLysTrpAsnAsnAlaAlaAspValValCysLys 301  
Db 778 CAATGGAGAGGCTGTGTGCTCCACATGGAGTCTGGCCAAATGTCATCTGTCTGT 837  
Qy 302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321  
Db 838 CAGCTCGGCTGGGAGTGGCATTCACCCCGGAGGAGGACACATCTGGTGGAGAAGGT 897  
Qy 322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341  
Db 898 GATCAGATCCTTAACAGCCGATTTACATGCTCTGGGGTGTGATCTCTCTGTGGATTGT 957  
Qy 342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361  
Db 958 COTGTGACTGCCCTGGGTGGTCTGACTGTCCCATGTCACAGCCCTCTGTGATCTGC 1017  
Qy 362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381  
Db 1018 TCAGAAAC-----1026  
Qy 382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401  
Db 1027 ---CAGATCCAGGTGCTTCCCGAG-----1047  
Qy 402 GlnAlaLeuValValCysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421  
Db 1048 -----TCAACAGCATCCGTG-----1062  
Qy 422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441  
Db 1063 TCTCAACCTACAGGCTCTGCGGCCTCA-----1089  
Qy 442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461  
Db 1089 -----1089  
Qy 462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481  
Db 1090 GAGGACAGCGCCCTACTGTCTCAGACAGCAGG-----CAGCTCCGCTGTGTGACCGG 1143  
Qy 482 HisSerProCysTyrGlyArgLeuValLysTyrGlnGlyGluTrpGlyThrValCys 501  
Db 1144 GCGGCTCCTCGCGGAGAGTGGAGATCTTGAACAGGCTCTCTGGGCGCCATCTGT 1203  
Qy 502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521  
Db 1203 -----1203

Db 1204 GATGACGCTGGGACCTGGACGATGCCCGGTGTGTGTGTCAGGACGCTGGGTGTGGAGAA 1263  
Qy 522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541  
Db 1264 GCCCTCAATGCCAGGGGTCTGCTCACCTTGGGGCAGGATCAGGCCCATCTGGTTGGAC 1323  
Qy 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561  
Db 1324 AACTTGAACCTCAGACGAAAGAGTCCACGCTGTGTGAGGTGCCCTTCCCGGGCTGGGG 1383  
Qy 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581  
Db 1384 CAGCACACTGCAGACACAGCAGCAGCGGGGGTTCATCTGCTCA-----GAGTCTCTG 1437  
Qy 582 GlyLeuArgLeuValGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSer 601  
Db 1438 GCCCTCAGGATGGTGTGAGTGGAGCAGCAGTGTGTGGTGTGCTGGAAGTTTCTACAA 1497  
Qy 602 GlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCys 621  
Db 1498 GGGACCTGGGCGAGTGTCTGCCGTAAACCCCATGGAAGACATCAGTGTGTCCACGATCTGC 1557  
Qy 622 SerGlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThr--- 640  
Db 1558 AGACAGCTGTGGTGTGGGACAGT-----GGACCTCACTTCTCTGTGCT 1605  
Qy 641 -----GlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSer 657  
Db 1606 CTTAGAGAAGTTTATAGGCCACAGTGGTGTGATAGATCCAGTGTGGAAAACTCACACC 1665  
Qy 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677  
Db 1666 TCTCTCTGGCAGTGTCTGACCTTGGAAATTACAACTCATGCTCTCCAAAGGAGGAA 1725  
Qy 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSer 697  
Db 1726 GCCTATATCTGTGTGCAGACAGC-----AGACAGATCCGCTGTGGATGGAGTGT 1779  
Qy 698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717  
Db 1780 CGCTGCTCTGGAGAGTGGAGATCTTACCAGGGCTCTCTGGGGCACCATCTGTGATGAC 1839  
Qy 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGlyCysGlySerAlaIle 737  
Db 1840 CGCTGGGACCTGGAGGATGCCGTGTGTGTGCAAGCAGCTGGGCTGTGGAGAAGCCCTG 1899  
Qy 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757  
Db 1900 GAGCCACTGTCTCTCTCTTCTGGGACGGATCAGGGCCCATCTGCTGGATCAAGTG 1959  
Qy 758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777  
Db 1960 AACTGCAGAGAGGAGGATCCCAAGTATGGAGGTGCCCTTCTGGGGATGGCGGCAACAC 2019  
Qy 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797  
Db 2020 AACTGCAATCATCAAGAGATGAGGAGTATCTGCTCAGGATTT-----GTGGCTGTG 2073  
Qy 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817  
Db 2074 GCTGGAGAGATGAGCCCTGCTCAGGGCGAGTAGAAGTGCATTCTGGAGAAGCCCTGGACC 2133  
Qy 818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837  
Db 2134 CCAGTGTCTGTGAAGAACTTCACACTCCCCACTGCCAGGTCATCTGTGAGAGCTGGGA 2193  
Qy 838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThr 857  
Db 2194 TGTGGCAAGGCTGTGTCTGCTCTGGGACACATGCCATTCAGAGAGTCCGATGGCCAGTGC 2253  
Qy 858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877  
Db 2254 TGGGCTGAAGAGTTACAGTGTGTGATGGGGGGGAGCGCTGAGCTCTGTGCTCTGCCCAAGTGT 2313

Qy	1337	gluThrTrpIleThrCysGlu	-----	1243
Db	3388	GAAGCCTACATCTCATGTGAAGGAAGAACCAAGAGCTGTCCAACCTGCTGCCGCGCTGC	3447	
Qy	1344	-----AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu	1260	
Db	3448	ACAGACAGAGAAGAGTCCGCTTCAGGGGAGGAGACACGAGTGTCTCAGGCGGGTGGAG	3507	
Qy	1361	IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla	1280	
Db	3508	GTGTGGCAACAACGGCTCTCTGGGCACCGTGTGCGATGACTCTCTGGAGCCTGGCAGAGGCT	3567	
Qy	1381	GluValValCysGlnLeuGlyCysGlySerAlaLeuAlaLeuAlaLeuArgAspAlaser	1300	
Db	3568	GAGTGTGTGTCTCAGCAGCTGGGCTGTGCCAGGCCCTGGAAGCCGTGCGGTCTGCAGCA	3627	
Qy	1301	PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysGlyGlyAsnGluSer	1320	
Db	3628	TTTGGCCCTCGAAATGGGAGCATCTGGCTGGAGAGGTGTCAGTGGCGGGGGCGGAGTCC	3687	
Qy	1321	PheLeuTrpAspCysHisAlaLysProTTPGlyGlnSerAspCysGlyHisLysGluAsp	1340	
Db	3688	TCCTGTGGGACATGTGTTGGGAGACCTCGGGGCAGAGGCTGCAACGACGAGGAGAT	3747	
Qy	1341	AlaGlyValArgCysSerGly	-----	1347
Db	3748	GCTGTGTGAGGTGCTGTGGTGAAGCAACATGCCCCACGACACAGCAGGGACACAGA	3807	
Qy	1348	-----GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuIleLeu	1364	
Db	3808	ACAACTTCAAAATCTCTCCCTGGCATCTTCTCCCTGCCTGGGGTTCTCTGCCTTATCCTG	3867	
Qy	1365	SerSerIlePheGlyLeuLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArg	1384	
Db	3868	GGTCCGCTCTCTCTCTGGTCTCTGTCGTCACCTCTGGTGACTCAGTACTCAGATGG	3924	
Qy	1385	ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu	1404	
Db	3925	GCAGAGGCG-----AGAGCCTTATCCAGCTATGAAGATGCTCTTGCT	3966	
Qy	1405	GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg	1424	
Db	3967	GAAGCTGTGTATGAGGAGCTCGATTACCTTCTGCACACAGAAGGAGGTCTGGGCAGCCCA	4026	
Qy	1425	ThrSer---AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu	1443	
Db	4027	GATCAGATGACTGATGCTCCCTGTAGAAATTTATGATGATGCTCTGAGAGTACCAGTGCT	4086	
Qy	1444	GlyValLeuProAlaSerGluAlaThrLys	1453	
Db	4087	GGAACTCCTTCTCCCTCTCAGGGGAATGAG	4116	
RESULT 7				
US-09-917-800A-1712				
; Sequence 1712, Application US/09917800A				
; Patent No. US20020119462A1				
GENERAL INFORMATION:				
; APPLICANT: Mendrick, Donna				
; APPLICANT: Porter, Mark				
; APPLICANT: Johnson, Kory				
; APPLICANT: Castile, Arthur				
; APPLICANT: Elashoff, Michael				
; APPLICANT: Gene Logic, Inc.				
TITLE OF INVENTION: Molecular Toxicology Modeling				
FILE REFERENCE: 44921-5038-US				
CURRENT APPLICATION NUMBER: US/09/917,800A				
CURRENT FILING DATE: 2001-07-31				
PRIOR APPLICATION NUMBER: US 60/222,040				
PRIOR FILING DATE: 2000-07-31				
PRIOR APPLICATION NUMBER: US 60/222,880				
PRIOR FILING DATE: 2000-11-02				
PRIOR APPLICATION NUMBER: US 60/290,029				
PRIOR FILING DATE: 2001-05-11				

```

RESULT 7
US-09/917-800A-1712
: Sequence 1712, Application US/09917800A
: Patent No. US20020119462A1
: GENERAL INFORMATION:
: APPLICANT: Mendrick, Donna
: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular
: FILE REFERENCE: 44921-5038-US
: CURRENT APPLICATION NUMBER: US/09/917-800A
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/290,029
: PRIOR FILING DATE: 2001-05-11

```

PRIOR APPLICATION NUMBER: US 60/290,645  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: US 60/292,336  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: US 60/295,798  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 60/297,457  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,884  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/303,459  
PRIOR FILING DATE: 2001-07-09  
NUMBER OF SEQ ID NOS: 1740  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1712  
LENGTH: 4344  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NW\_022849  
US-09-917-800A-1712

Alignment Scores:  
Pred. No.: 178e-105 Length: 4344  
Score: 1142.50 Matches: 331  
Percent Similarity: 41.73% Conservative: 136  
Best Local Similarity: 29.58% Mismatches: 343  
Query Match: 14.04% Indels: 312  
DB: 10 Gaps: 39

US-09-759-130B-381 (1-1453) x US-09-917-800A-1712 (1-4344)

QY 258 SerAspLeuGluValArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeu 277  
Db 10 TCTGTTTGGCTGAGGCTGGTGAATGAGGAGACAGGTGCGGGTGGGAGATC 69  
QY 278 LysIleGlnGlyArgTrpGlyThrValCysHisLysTrpAsnAsnAlaAlaAsp 297  
Db 70 CTTTACAGGGTCTCTGGGGCCACCATGTGTATGACAGCTGGGACATCAATGATGCCAAC 129  
QY 298 ValValCysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeu 317  
Db 130 GTGGTGTGAGGAGCTGGGCTGTGGCTGTGGCTGTCTGCCCGAGAGTCCCGAGTTT 189  
QY 318 GlnSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPhe 337  
Db 190 GGACAGGGTCTGGTCCCATTTCTGGATGACGTGGCTGAGGAGACATGAGGCTAT 249  
QY 338 LeuTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspVal 357  
Db 250 CTGTGGAGCTGTCTCCACCGAGGCTGGCTCTCTCAATACTGTGGACATCAGGAGGATGCT 309  
QY 358 SerValIleCysSerAspGlyAlaAspLeuGluValArgLeuAlaAspGlySerAsnAsn 377  
Db 310 GGAGTGTATCTGTACAT-----TCTCAACACAGCATGCC 345  
QY 378 CysSerGlyArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsn 397  
Db 346 ACACCCGGT-----TGTTGG-----AACCCCGGG 369  
QY 398 TrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPhe 417  
Db 370 GGCACAAATAACGATGTGATCTACGACACTCAAGAAACC----- 408  
QY 418 GlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSer 437  
Db 409 ---ACAGAACTTCTCAACACAGCATGCCACACCTGATGGTGG----- 450  
QY 438 CysThrGlyAsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThr 457  
Db 451 -----AACCATGGGGGCACAAATTAATGATGTGATCTATGAC-----ACTCAAGAA 495  
QY 458 CysPheArgArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArg 477

Db 496 ACCACAGAGAACACAGATTCTGGT-----TTGGCTGTGAGG 531  
QY 478 LeuValGlyAlaHisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrp 497  
Db 532 CTGGTGAATGAGGAGACAGGTGTCGGGGTGTGTGGAGATCCTTTACAGGGTTCCTGG 591  
QY 498 GlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeu 517  
Db 592 GGCACCGTGTGTGACGACAGCTGGGACATCAATGATGCCAAGCTGGTGTGACGAGCTG 651  
QY 518 GlyCysGlyLysProMetHisValPheGlyMetThrTrpPheLysGluAlaSerGlyPro 537  
Db 652 GGCTGTGGCTGGGCTTGTCTGCCCGAGGAGTGGCCAGTTTGGACAGGGCTCTGGGTCC 711  
QY 538 IleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHis 557  
Db 712 ATTGTCTGGATGACGTGGCTGTAGAGGACATGAGGCTATCTCTGGAGCTGCTCCAC 771  
QY 558 SerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer--- 576  
Db 772 CGAGGCTGGCTCTCTCAATACTGTGACATCAGGAGGATGCTGGAGTCTATCTGTTCATAT 831  
QY 576 ----- 576  
Db 832 TCTCAACACAGCAGTCCACACCCCGATTCTCAACACAGCAGTCCACACCCCGTTGGTGG 891  
QY 577 -----GlyAspAlaThrTrpGly----- 582  
Db 892 AACCCCGGGGGCACAAATAACGATGTGTCTATGGACCGGACACAGACACAGCGCAACA 951  
QY 583 -----LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGlu 597  
Db 952 GATTCGTGGTTCGGCTGTAGGCTGGTGAATGGAGGAGACAGGTGTCAGGCTCGTGGAG 1011  
QY 598 ValTyPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAla 617  
Db 1012 ATCTTTACAGGGTTCCTGGGGTACCGTGTGTGACGACAGCTGGGACACAGGATGCC 1071  
QY 618 AlaValValCysSerGlnLeuAspCysProSerIleIleGlyMetGlyLeuGlyAsn 637  
Db 1072 AACGTGGTGTGACGAGCAGCTGTGTGGCTGGCTGTCTGCCCGAGAGTGGCCAC 1131  
QY 638 AlaSerThrGlyTyGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSer 657  
Db 1132 TTTGGACAAAGGCTCTGGATCCATTGTTCTGGATGACGTGGCTGTACAGGACATGAGGC 1191  
QY 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAspCysSerHisSerGluAsp 677  
Db 1192 TATCTGTGGAGCTGTCCCGAGGCTGGCTCTCTCAATACTGTGGCCACCATGAGGAT 1251  
QY 678 ValGlyValIleCysSerAspAlaSer-----AspMet----- 688  
Db 1252 GCTGGAGTCTATGTTTCAGATGCCCAACCCAGACACACCTGGCCAGATATGTGGCT 1311  
QY 689 -----Glu-LeuArgLeuValGlySerSerArgCysAlaGlyLysValG 704  
Db 1312 ACTACCACTCCAGAACTACAACAGATTGTGGAGT-ACAAAATATCTCTCTCTCTCC 1370  
QY 704 vAlaAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAl 724  
Db 1371 TACAACACAAATTCGCCACCATTA-----GCCGATTGGTGACAACTCTTCTCC 1418  
QY 724 aGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer----- 740  
Db 1419 GGAATACACC-----TGTTGAGGTTTACTGACCCCTACCCCTATGGCAGTT 1463  
QY 741 -ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerClyCysTh 760  
Db 1464 TTCCAGCCCATCATCCCTGGAGC-----TATCCTTAACATGCCAGATG--- 1509  
QY 760 rGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysH 780

Db 1510 -----TTGTGG----- 1515  
Qy 780 sLeuAaMetGluAlaSerLeuLeuCysSerAlaHisArgGlnProArgLeuValGlyAl 800  
Db 1516 -----AAATTTTCGTCTCCAGCATCAACCGTGTACAGTGGTCTTCAC 1559  
Qy 800 aAPMetProCysSerArgValGluValLysHisAlaAspThrTrpArgSerValCy 820  
Db 1560 AGATGTGCAGCTTGAAGAGGT-----TG 1583  
Qy 820 sAPSerAspPheSerLeu-----HisAlaAlaAsnValLeuCysAr 834  
Db 1584 CAATATGACTACATCTCGTTTGTGAGCCCTGAAACAATCTCTCTCATTTGTCG 1643  
Qy 834 gGluLeuAaCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGly-- 853  
Db 1644 G-----GTTTGTGATGGGTTCATGATCTTT 1670  
Qy 854 -----AsnGlyLeuThrTrpAlaGl 860  
Db 1671 CACCTCAACCCAGAACTTTCATGCTGTAGTCTTTATCACGGATGGCAGTGTACAGGAG 1730  
Qy 860 uLysPheGlnCysGlu-----GlySerGluThrHisLeuAlaLe 873  
Db 1731 AGGTTTCCAAAGCTGACTACTTCCACTATCAGCAGCAGCAACCTCTCCAAACGAC 1790  
Qy 873 uCysProIleVal-----GlnHisProGluAspThrCys----- 1885  
Db 1791 GTTCCCGATCGTTACTGATGGTGACAACTCTCTCCGGAATACACCTGTGGAGGTTT 1850  
Qy 885 eHisSerArgGluValGlyValValCysSerArgTrpThrAspValArgLeuValAsnGl 905  
Db 1851 ACTGACCTACCTATGGCAGTGTTCAGCCCATAC----- 1887  
Qy 905 yLysSerGlnCysaspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCy 925  
Db 1888 -----TACCCTGGAGACTAT- 1902  
Qy 925 sAspThrHisTrpAspProGluAspAlaArgValLeuCysArg----- 939  
Db 1903 -----CCTAACATGCCAGATGTTGTGGAAATTTTCGTPCCCGACAT 1946  
Qy 940 -----GlnSerCysGlyThr----- 945  
Db 1947 GAACCGTGTGACAGTGTCTTCACAGATGTCAGCTTGAAGAGGTGGCAACTATGACTA 2006  
Qy 946 -AlaLeuSerThrThrGlyGlyLysTyrlleGlyGluArgSerValArgValTrpGlyHi 965  
Db 2007 CATCTCGGTTTGTGATGCTGCTGAATACAATCTTCTCTCATCTCTCGGGT----- 2058  
Qy 965 sArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGl 985  
Db 2059 -----TGTGATGGTCCATGATCTTTCACCTCAACCCAG----- 2094  
Qy 985 yAlaProCysIleHisGlyAsnThrValSerVal-----IleCysThrGlySerLe 1003  
Db 2095 -----AACTTCATGCTGTAGTCTTTATCACGGATGGCAGTGT 2132  
Qy 1003 uThrGlnProLeuPhe-----ProCysLeuAlaAsnValSerAs 1016  
Db 2133 CACGAGGAGGGTTCACAGCTGACTACTTCCACTCTTATCAGGACCAACCACTCC 2192  
Qy 1016 pProTyTrLeuSerAlaValProGluGly-----SerAlaLeuIleCysLeuGluAspLy 1034  
Db 2193 TCCACACGCTTCCCGATCATTAAGTGAATGATCTTCATGTGG----- 2238  
Qy 1034 sArgLeuArgLeuValAspGlyAspSerArgCysAlaGArgValGluIleTyHisAs 1054  
Db 2239 -----CTGAGGCTGTAATGAACAAACCGGTGTGAGGCGGAGTGGAGATCTTGTACAG 2294  
Qy 1054 pGlyPheTrpCpLyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCy 1074  
Db 2295 AGGCTCTTGGGTACCGTGTGCCGACGACGTGGGACATCAATGATGCCAATGGTGTCTG 2354

Qy 1074 sGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGl 1094  
Db 2355 CAGACAGCTCGTGTGGCTGCTGCTCTGCTGCCAGAAATGCTTGTGTGGTCAGGG 2414  
Qy 1094 ySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGl 1114  
Db 2415 TTCAAGGCTCATTCCTCGTGGATGATGTCTGTCTGTGGTATGAGTCCACCTGTGAA 2474  
Qy 1114 nCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyVal 1134  
Db 2475 TTGTCGTCAACCTCGCTGGCTTGTTCATAATGTGCTCATGTTGAGGATGACGAGGATCAT 2534  
Qy 1134 eCysSer-----GluPheThrAlaLeuArgLeuTyTrpGluThrGluTrpGluSerCy 1152  
Db 2535 TTGTCCTACTCCCTCATCCGACTCC-CTCTCCTGGTCCAGTTTGGACAAGTCTCTCTTTTG 2593  
Qy 1152 sAlaGlyArgLeuGluValPheTyTrpAsnGlyThrTrpGlySerValGlyArgAsnIl 1172  
Db 2594 T-----AACTATACTTGTGGAGTTTCTGACTGGAC-TCTCTGGCAATTTTCTAGC 2646  
Qy 1172 eThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValVa 1192  
Db 2647 CCATCTACCTCGGAGCTATCTTAATAATGCCAGATTTTGTGGA----- 2692  
Qy 1192 iSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp----- 1209  
Db 2693 -----ACATTGAAGTCCCAACACTACCGCTGACTGTGGCTTCCAGAGATGTGCAG 2745  
Qy 1210 -----IleGlnCysProLysThr-----Hl 1216  
Db 2746 CTGAAGGGGGCTGCAACTATGACTATATAGAGATTTTGTATGGCCCCACACAGTTCA 2805  
Qy 1216 sIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro 1234  
Db 2806 CCTCTCATGTCGGGTTTGTATGGGCCCATGAGGCTCTTTCACCTTCAACATCCA 2860

RESULT 8

US-09-147-947-5  
; Sequence 5, Application US/09147947A  
; Patent No. US20020160490A1  
; GENERAL INFORMATION:  
; APPLICANT: TSURUOKA, No. US20020160490A1uo  
; APPLICANT: YAMASHIRO, Kyoto  
; APPLICANT: YAMAGUCHI, No. US20020160490A1om1  
; TITLE OF INVENTION: No. US20020160490A1el Serine Protease  
; FILE REFERENCE: 001560-349  
; CURRENT APPLICATION NUMBER: US/09/147,947A  
; EARLIER FILING DATE: 1997-03-24  
; EARLIER APPLICATION NUMBER: PCT/JP98/03324  
; EARLIER FILING DATE: 1998-07-24  
; EARLIER APPLICATION NUMBER: JP 9/213969  
; EARLIER FILING DATE: 1997-07-24  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 2562  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; OTHER INFORMATION:  
US-09-147-947-5

Alignment Scores:  
Pred. No.: 4,26e-74 Length: 2562  
Score: 830.00 Matches: 260  
Percent Similarity: 40.26% Conservative: 106  
Best Local Similarity: 28.60% Mismatches: 349  
Query Match: 10.20% Indels: 195  
Gaps: 34

US-09-759-130B-381 (1-1453) x US-09-147-947-5 (1-2562)



Db 1863 AGCACACTGTTTCAAGAGGTATGGCAACAGCACTAGCAGCTATGCTGTTAGGGTTGGAGA 1922  
QY 848 pHis-----PheGlyLysGlyAsnGly-----LeuThrTr 858  
Db 1923 TTATCATACTCTGGTACCAGAGAGTTTGGAGAGAAATGGAGTTCAACAGATTGTGAT 1982  
QY 858 pAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValG1 878  
Db 1983 TCATCGGGAGTATCGACCGCAGCGAGTATTATGACATAGCCCTGTTAGATTA---CA 2039  
QY 878 nHisProGluAspThrCysIleHis---SerArgGluValGlyValValCysSerArgTy 897  
Db 2040 AGGACCAAGAGAGCAATGTCAGATTCACGACGCCATGTTTCCACGCCCTTTACCACT 2099  
QY 897 rThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnVa 917  
Db 2100 CTGAGAGAGAGCCACAGAAACAGCATCCACTGT-----TA 2138  
QY 917 lLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLe 937  
Db 2139 CATAACAGGATGGGT-----GACACAGGAGCAGGCCTA 2171  
QY 937 uCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrlleGlyG1 957  
Db 2172 TTCAGAACACTACAA---CAACAGCAGCATTCCTTACTTCCATAAAGGTTTGTGAAGA 2228  
QY 957 uArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsn-----GluSe 974  
Db 2229 ACCTTATAAGGTCGGTTTACAGGAGGAGTCTTGTGCTGGAACCTCCATGAACACAA 2288  
QY 974 rLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaPro-ProCysIleHisGlyAsnT 994  
Db 2289 ACAGTGGACAGCTGCCAGGAGAC---AGCGGAGGACCACCTCATGTCGTGAACGCCCGG 2345  
QY 994 hrValSerValIleCysThrGly 1001  
Db 2346 AGAGAGCTGGGTGTGATGGG- 2368

## RESULT 9

US-09-918-995-32438 20030073623

Sequence 32438, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

FROM VARIOUS CDNA LIBRARIES

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

PRIOR FILING DATE: 2001-07-30

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 32438

LENGTH: 473

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(473)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-32438

Alignment Scores:  
Pred. No.: 6.13e-74 Length: 473  
Score: 818.00 Matches: 143  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.05% Indels: 0  
DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x US-09-918-995-32438 (1-473)

QY 1194 LeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysPro 1213  
Db 42 CTCGCCCTTTATCTAAGACAGAGCTCTGGTTTCATGTGGTGATGACATTCAGTGTCTCT 101  
QY 1214 LysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSer 1233  
Db 102 AAAACGCATATCTCCATATGCGAGTGCCTGTCTGCCCATGGGAGCGAAGAAATCTCCAGC 161  
QY 1234 ProAlaGluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThr 1253  
Db 162 CCAGCAGAGAGACCTGGATCACAATGGAAGATAGATAGAGTCGCTGGAGGAGACACC 221  
QY 1254 GluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThrValCysAspAsp 1273  
Db 222 GAGTGTCTCTGGAGAGTGGAGATCTGCCACGACGCTCTGGGCGCAGTGTGTGTGATGAC 281  
QY 1274 SerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeu 1293  
Db 282 TCTTGGGACCTGCCGAGCGGAGTGTGTGTGAGCAGCTGTGGCTGTGGCTCTGCTCTG 341  
QY 1294 AlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMet 1313  
Db 342 GCTGCCCTCAGGAGCGCTTCGTTGGCCAGGAACTGGAACCATCTGTTGGATGACATG 401  
QY 1314 ArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSer 1333  
Db 402 CGGTGCAAGGAATGAGTCAATTTCTATGGACTGTTCACGCCAAACCTGGGGACAGAGT 461  
QY 1334 AspCysGly 1336  
Db 462 GACTGTGGA 470

## RESULT 10

US-09-147-947-3

Sequence 3, Application US/09147947A

Patent No. US20020160490A1

GENERAL INFORMATION:

APPLICANT: TSURUOKA, No. US20020160490A1uo

APPLICANT: YAMASHIRO, Kyoto

APPLICANT: YAMAGUCHI, No. US20020160490A1om1

TITLE OF INVENTION: No. US20020160490A1el Serine Protease

FILE REFERENCE: 001560-349

CURRENT APPLICATION NUMBER: US/09/147,947A

CURRENT FILING DATE: 1997-03-24

EARLIER FILING DATE: 1998-07-24

EARLIER FILING DATE: 1997-07-24

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 2614

TYPE: DNA

ORGANISM: Mouse

FEATURE:

OTHER INFORMATION:

US-09-147-947-3

Alignment Scores:  
Pred. No.: 3.99e-63 Length: 2614  
Score: 722.00 Matches: 207  
Percent Similarity: 41.54% Conservative: 85  
Best Local Similarity: 29.45% Mismatches: 260  
Query Match: 8.87% Indels: 153  
DB: 9 Gaps: 25

US-09-759-130B-381 (1-1453) x US-09-147-947-3 (1-2614)

QY 10 IleAspPheGlyArgCysCysHisGlnAsnLeuPheSerAlaValThrCysIle 29  
Db 692 GTAGACTGGGGCTACTCGCATTTGTCAAGCGCCCGCGTGTGCCGTCATT----- 742  
QY 30 LeuLeuAsnSerCysPheLeuIleSerPheAsnGlyThrAspLeuGluLeuArg 49

Db	743	-----		CGC	745
Qy	50	LeuValAsnGlyAspClyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrp	69		
Db	746	CTTGTGGTGGGAACAGTGGCGATGAAGTCGATGGAGCTGTACACAGCTGCCAGTGG	805		
Qy	70	GlyThrValCysAspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeu	89		
Db	806	GGACCATCTGTGACGACCAATGGGACAAATCAGACGCAGACGCTCACTGTAGGCAGCTG	865		
Qy	90	-----GlyCysProPheSerPheAlaMetPheArgPheGlyGlnAlaValThrArg	106		
Db	866	GGGCTCAGTGGCATGCCAAGCATGGCATCAGGCACATTTTGGGAAGA	916		
Qy	107	HisGlyLysIleTrpLeuAspValSerCysTyrglyAsnGluSerAlaLeuTrpGlu	126		
Db	917	TCGTGGCCCAATATGTGTGATGAAGTGCACCGGAAACAGAGCTGTCAATGTAGACA	976		
Qy	127	CysGlnHisArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsn	146		
Db	977	TGTCCAAAGAGTTCCTGGGGCCCAACATACCTGTGCCATAAAGAGATGCTGGAGTGTCT	1036		
Qy	147	CysTyrglyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGly	166		
Db	1037	TGTGTCTTCTAACAGATGGTCTCATCAGATGGCGAGGAGAAAGTACCATGAAGGT	1096		
Qy	167	ArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeu	186		
Db	1097	CGCCTGGAGTCTACTACAAAGGGGAGTGGGGACAGTCTCTGTATGATGGCTGGACTGAG	1156		
Qy	187	AsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerPheIleSerSerGly	206		
Db	1157	ATGACACATACGTGGCTGTGCAGTCTGGGATTTAATACGGCAACAGTCTCTCT	1213		
Qy	207	ValValAsnSerProAlaValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGly	226		
Db	1214	GTCAACCATTTTGTATGGCAGCAACAGGCCCATATGGCTGGATGAGTCAGTCAGTCTCAGGA	1273		
Qy	227	AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis	246		
Db	1274	AAAGAAGTCAGCTTCATTCAGTGTTCAGGAGACAGTGGGGAAGGCATGATGCAGCAT	1333		
Qy	247	AsnGluAspValThrLeuThrCysTyraSpSerSerAsp	259		
Db	1334	AGAGAAGATGGGCTTCACCTGCTATCCTGCACCGATGCACATAGGCCTTCTCCAGGT	1393		
Qy	260	LeuGluLeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuLysIle	279		
Db	1394	TTTCCCATCAGACTAGTGGATGGAGAGAATAAGAAGGAAGCAGTGGAGGTTTGTGTC	1453		
Qy	280	GlnGlyArgTrpGlyThrValCysHisLysTyTrpAsnAsnAlaAlaAlaAspValVal	299		
Db	1454	AATGGCCCAATGGGGAACAATCTCGCATGACGGATGGCCGATTAAGCATGATGCAGTCTCATC	1513		
Qy	300	CysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSer	319		
Db	1514	TGCGGCGAGCTTGGCTATAGGGTCTGCCAGACAGGACTATGGCTTATTTTGGGAA	1573		
Qy	320	GlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrp	339		
Db	1574	GGAAAGGCCCATCCATCGATGGATAATGTGAAGTCACAGGAATAGAAAGGCCCTGGCT	1633		
Qy	340	AspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerVal	359		
Db	1634	GACTGTGTCAAAACAAGACATTTGGAAGGCACAACTGCCCCACAGTGAAGTGCAGAGTCT	1693		
Qy	360	IleCys	372		
Db	1694	ATCTGTGACTATTTAGAGAAGAAAGCATCACTAGTGTGTAATAAGAGATGCTC	1747		
Qy	373	AspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGlu	390		

D	b	1748		- - - - TCAFTCTGGATGTGGACTGAGGTTCATTGCACCCCTCG- GCAGAACAAGCATTTGGT	1806
Q	y	391	T	rPThrIleCysAspGlnA <sup>s</sup> nTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeu	410
D	b	1801	TGG	:    :	
D	b	1801	TGG	- - - - GAACAATTCITTAAGGGTGCCCTGGCCTTGGCA-	1836
Q	y	411	G	lycCysProPheSerValPheGlySerArgArgAlaLysProSerAsnGluAlaArgasp	430
D	b	1837	GGCTTC	CCTCAGGCTCAGGCTCGGCCCATGGAGACGGCAGGCTGCT-	1881
Q	y	431	I	letrp-----lleasrSeriileSerCystHrGlyAsnGlu---	442
D	b	1882	TTTTGG	AAGCTACCCTTCTGAGTAGCTGCTGGGTCTCGACGCTGCACACTGCACACTCTCAAAG	1941
Q	y	443	-----	-----SerAlaLeuTrpAspCystHrTyrrAspGly--	452
D	b	1942	GTACGGAA	CAACTCGAGGAGCTATCCAGTTCGAGTTGGGATTATCATACTCTGGTACC	2001
Q	y	453	-----	-----lysAlaLysArgThrCysPheArgSerAspAlaGly	465
D	b	2002	AGAGGAGT	TTTGAACAATAAGGGTTCAACAGATTCT- -----GATTCAACAG	2049
Q	y	466	Vallile	cysSer-AspLysalaAspleuAspleuValGlyAlaHisSerProCy	485
D	b	2050	GAAC	TACAGCCAGCACAGAGGAGCTATGACATGCCCTGGTGTAGATTGCAAGAACCA---	2107
Q	y	485	sTYrGly	ArgLeuGluValLysTyrglnGlyltpGlyThrValCysHisAspArgTr	505
D	b	2108	----	GGGGAGCAATGTGCCAGACTAAGCACCCACGCTTTTGCCAGCGCTGTATACCTCTATG	2163
Q	y	505	pSerThr	Arg-----AsnAlaAlavalValCysLysGlnLeuGlyCysGlyLysPromE	523
D	b	2164	GAGAGAG	GCCACACAAAAACGCCCTCCAACGTGCACATAACAGAGTGGGGAGACACAGG	2223
Q	y	523	thisValphe	GlyMetHr-----TyrPh	531
D	b	2224	TCGTGCT	ACTCAAGACTCTACACAACGCTGCTGTGCTTWTCCCAGAGGTTTGT	2283
Q	y	531	eLysGlu	AlaserGlyproIletrPleuAspValserCysIleGlyAsnGluSerAs	551
D	b	2284	TAAAGAGG	TACAAGGGACTATTACTGGGAGAATGCTCTGCTGCTGGGAAAC-----	2335
Q	y	551	niletr	aspCysgluHisserclyTrpGlyLysHisasnCysvalHisArgrGluAspva	571
D	b	2336	-----	-----CTCCAAGAAGCAACACCG	2352
Q	y	571	llileValthr	CysSerGlyAspAlathrtrpGlyLeuArgLeuValGlyGlySerAsnar	591
D	b	2353	TGTGGAC	GTGCCAGCTGCCAGGAGACAGT-----GGAGGACCACTCAT	2391
Q	y	591	gCysSer	GlyArgLeuGlu-----ValTyrPheGlnGlyArgTrpglyThrValcy	608
D	b	2392	GTGTGAAA	AGGCTGATGAGTCCTGGGTTGTGTATGGGTGACTTCTCCGGGGTATGGATG	2451
Q	y	608	s-----	AspaspGlyTrpasnserLysAlaAlaValValCysSerGlnLeuas	625
D	b	2452	TGGAGT	CAAGACAC-TCCCTGGAGTTTATACAGAGTCCCCGCCCTTGTACC-----TTGGA	2507
Q	y	625	p-----	CysProSerSeriileglyMetGlyLeuGlyAsnAlaserThrGlytyrGl	643
D	b	2508	TAAAGTG	TCACCAGCTCTGTAAC-----TATGG	2537
Q	y	643	yLysile	645	
D	b	2538	AAAGCTC	2544	

RESULT 11  
US-09-905-291A-147  
; Sequence 147, Application US/09905291A  
; Patent No. US20020160374A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi

Alignment Scores:	
Pred. No.:	2.64e-60
Score:	691.50
Percent Similarity:	52.51%
Length:	1686
Matches:	148
Conservative:	40

Query Match:	41.34%	Mismatches:	139
DB:	9	Indels:	31
		Gaps:	9
US-09-759-130B-381 (1-1453) x US-09-905-291A-147 (1-1686)			
Qy 21	LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys-----Phe	36	
Db			
Db 107	CTATTCTCC-----TTGATCTTGCATTGTGCACGAGACTGGATT	148	
Qy 37	LeuIleSerPheAsnGlyThrAspLeuGluLeuValAsnGlyAspGlyPro	56	
Db			
Db 149	CTACGCTCTCCACTCGA-----GTGCGCTGGTGGGGGCTCCACGC	193	
Qy 57	CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspGly	76	
Db			
Db 194	TGTGAAGCGCGGTGGAGGTGGAACAGAACGCCAGTGGGGCCTGTGTGATGACGC	253	
Qy 77	TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla	96	
Db			
Db 254	TGGGACATTAAAGACGCTGGCTGTGTGTCGCCGGAGCTGGCTGTGGAGCTGCCAGCGA	313	
Qy 97	Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp	113	
Db			
Db 314	ACCCTAGTGGTATTGTATGAGCCAGCAGCAAGAAAGACAAAGTCTCTCAAC	373	
Qy 114	AspValSerCysTyrrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly	133	
Db			
Db 374	TCAGTCAAGTTCACAGACAGATACATTTGGCTCAGTGTGACCAAGAGAA-----	427	
Qy 134	SerHisAsnCTyrHisGlyGluAspValGlyValAsnCys-----	147	
Db			
Db 428	GTATTATGTTTCACATGATGAAGATGCTGGGGCATCTGTGAGAACCCAGAGAGCTCT	487	
Qy 148	TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg	167	
Db			
Db 488	TTCTCCCATCCAGAGGTGTCAGGCTGGCTGAGCGCCCTGGGCATTGCAAGGGAGCG	547	
Qy 168	ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspGlyTrpAsnLeuAsn	187	
Db			
Db 548	GTGGAAGTGAAGCACCAGACCACTGATACCGTGTGCCACAGAGCTGGAGCTCCGG	607	
Qy 188	ThrAlaAlaValValCysArgGlnLeuGlyCysProSerPheIleSerSerGlyVal	207	
Db			
Db 608	GCCGAAAGTGTGTGCCGCGACCTGGATGTGGGAGGCTGTACTGACTCAAAAACGC	667	
Qy 208	ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspIleLeuCysGlnGly	226	
Db			
Db 668	TGCAACAAAGCATGCCATGCCCCAAACCATCTGGCTGAGCCAGATGCTATGCTCAGGA	727	
Qy 227	AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis	246	
Db			
Db 728	CGAGAACCAACCTCTCAGATTCCTCTGGGCTTGGGGAGAACCACTTGAACCAT	787	
Qy 247	AsnGluAspValThrLeuThrCysTyrrAspSerSerAspLeuGluLeuArgLeuValGly	266	
Db			
Db 788	GATGAACACACGTGGGTGCGAATGTGAAGATCCCTTTGAC-----TTGAGACTAGTAGA	841	
Qy 267	GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal	286	
Db			
Db 842	GSAGACAACCTCTGCTCTGGGCGACTTGGAGTGTGTCACAAAGGGCGTATGGGGCTCTGTC	901	
Qy 287	CysHisIleIleTyrrAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGly	306	
Db			
Db 902	TGTGATGACAACCTGGGGAGAAAGGAGGACCAAGTGGTATGCAACCACTGGGCTGTGGG	961	
Qy 307	ThrAlaLeu-----HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal	324	
Db			
Db 962	AAGTCCCTCTCCCTCTCAGACACCGGAATGCTATGGCCCTGGGGTGGCGGCATC	1021	
Qy 325	TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer	344	
Db			
Db 1022	TGGCTGATTAATCTGCTGTGTCAGGGAGGAGCAGTCCCTGGAGCACTGCCAGCAGACA	1081	

Qy 345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362  
Db 1082 TTTTGGGGTTCACGACTGCACCCAGGAGATGTGGTGTCTCATCTGCTCA 1135

## RESULT 12

US-09-902-853-147

; Sequence 147, Application US/09902853

; Publication No. US20020192659A1

## ; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth, J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/902,853

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: US/09/665,350

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

; PRIOR FILING DATE: 1999-11-29

; PRIOR APPLICATION NUMBER: PCT/US99/28313

; PRIOR FILING DATE: 1999-11-30

; PRIOR APPLICATION NUMBER: PCT/US99/28564

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/28565

; PRIOR FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: PCT/US99/30095

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: PCT/US99/30911

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US99/30999

; PRIOR FILING DATE: 1999-12-20

; PRIOR APPLICATION NUMBER: PCT/US00/00219

; PRIOR FILING DATE: 2000-01-05

; NUMBER OF SEQ ID NOS: 423

; SEQ ID NO 147

; LENGTH: 1686

; TYPE: DNA

; ORGANISM: Homo Sapien

US-09-902-853-147

## Alignment Scores:

Pred. No.: 2,64e-60 Length: 1686

Score: 691.50 Matches: 148

Percent Similarity: 52.51% Conservative: 40

Best Local Similarity: 41.34% Mismatches: 139

Query Match: 8.50% Indels: 31

DB: 9 Gaps: 9

US-09-759-130b-381 (1-1453) x US-09-902-853-147 (1-1686)

Qy 21 LeuPheSerAlaValThrCysIleLeuLeuLeuAsnSerCys-----Phe 36  
Db 107 CTATTCTCC-----TTGATCCTTGCATTGCACACAGACTGGATTC 148  
Qy 37 LeuIleSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56  
Db 149 CTAGCCTCTCCATCTGGA-----GTGCGCTGGTGGGGGCTCCACCGC 193  
Qy 57 CysSerGlyThrValGluValLysPheGlnGlyInTrpGlyThrValCysAspGly 76  
Db 194 TGTGAAGGGCGGTGGAGTGGAGAACAGAGCCAGTGGGGCAGCTGTGTGATCAGCGC 253  
Qy 77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96  
Db 254 TGGGACATTAAAGGACGTGGCTGTGTGGCGGAGCTGGGCTGTGGAGCTCCACGCGA 313  
Qy 97 Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113  
Db 314 ACCCTAGTGTATTTTGTATGAGCCACAGCAGAGAAAGACCAAGGTCTCATCCAA 373  
Qy 114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133  
Db 374 TCAGTCAAGTTCACAGAGAACAGAGATACATTGGCTGAGTGTGAGCAAGAAGAA----- 427  
Qy 134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147  
Db 428 GTTATGATTGTTCACATGATGAAGATGCTGGGGCATGCTGTGAGAACCCAGAGACTCT 487  
Qy 148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167  
Db 488 TTCTCCCCAGTCCAGAGGGGTGTGAGGCTGCTGAGCGCCCTGGGCTTGAAGGACGC 547  
Qy 168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspGlyTrpAsnLeuAsn 187  
Db 548 GTGGAAGTGAAGCACCAGAACCATGTTATACCGTGTGCCAGACAGGTGGAGCTCCGG 607  
Qy 188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerGlyVal 207  
Db 608 GCCGCAAGGTGTGTGCGGCGAGCTGGAGTGGAGGGCTGTACTGTACTCAAAACGC 667  
Qy 208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226  
Db 668 TGCAACAAGCATGCCTATATGCGCGAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 727  
Qy 227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246  
Db 728 CGAAGCAACCTTTCAGGATTCGCTTGTGGGCGCTTGGGGGAGAACACCTGCAACCAT 787  
Qy 247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuLeuArgLeuValGly 266  
Db 788 GATGAAGACACGTGGTCAATGTGAAGATCCCTTTGAC-----TTGAGACTAGTAGGA 841  
Qy 267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286  
Db 842 GGAGACAACCTCTGCTCTGGCGGAGTGGAGGTGCTCCACAGGCGCTATGGGCTGTGTC 901

```
QY 287 CysHisLysTrpAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
   |||
   |||
Db 902 TGTGATCACAACGTGGGAGAAAAGGAGGACCGTGTATGCAAGCAACGTGGGTGTGGG 961
   |||
QY 307 ThrAlaLeu-----HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
   |||
Db 962 AAGTCCCTCTCTCCCTCTCAGAGACCGGAAATGCTATGCGCCCTGGGGTGGCCGCATC 1021
   |||
QY 325 TrpLeuAspGlyValSerCysSerGlyAsnGlnSerPheLeuTrpAspCysArgHisSer 344
   |||
Db 1022 TGGCTGATGATGTTCTCTCAGGAGAGAGAGAGTCCCTGGAGCAGTCCACACAGAGA 1081
   |||
QY 345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
   |||
Db 1082 TTTTGGGGGTTTCAGCAGTCCACCCAGGAGAGATGCTGGTGTCTCTCTCTCA 1135
   |||

RESULT 13
US-09-907-824-147
; Sequence 147, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Flivaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Guney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
```

```
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 147
; LENGTH: 1686
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-824-147

Alignment Scores:
Pred. No.: 2.64e-60 Length: 1686
Score: 691.50 Matches: 148
Percent Similarity: 52.51% Conservative: 40
Best Local Similarity: 41.34% Mismatches: 139
Query Match: 8.50% Indels: 31
DB: 9 Gaps: 9

US-09-759-130B-381 (1-1453) x US-09-907-824-147 (1-1686)
QY 21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCys-----Phe 36
   |||
   |||
Db 107 CTATTCCTCC-----TTGATCTTGCCATTGACACGACCTGGATTC 148
   |||

QY 37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
   |||
   |||
Db 149 CTAGCGTCTCCATCTGGA-----GTGCGCTGGTGGGGGCTCCACCGC 193
   |||

QY 57 CysSerGlyThrValGluValLysPheGlnGlyThrValGlyThrValCysAspGly 76
   |||
   |||
Db 194 TGTGAAGGGCGGTGGAGGTGGAACAGAAAGGCCAGTGGGCGACGCTGTGTATGACCGC 253
   |||

QY 77 TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
   |||
   |||
Db 254 TGGACATTAGAGCGTGGCTGTGTGTGCGGGAGCTGGCTGTGGAGTCCACGCGA 313
   |||

QY 97 Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
   |||
   |||
Db 314 ACCCTAGTGTATTTTGTATGAGCCACGACAGAAAAAGAGAGAAAGTCTCATCCA 373
   |||

QY 114 AspValSerCysTyrGlyAsnGlnSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
   |||
   |||
Db 374 TCAGTCAGTTGCACAGGAACAGAAATACATTGGCTGCTAGTGTAGCAAGAAAGAA 427
   |||

QY 134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
   |||
   |||
Db 428 GTTTATGATTTGTCATGATGAGATGCTGGGCGATCTGTGAGAACCACAGAGCTCT 487
   |||

QY 148 TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
   |||
   |||
Db 488 TTCTCCCGCAGTCCACAGAGGGTGTGAGGCTGGCTGACGCGCCCTGGGCATTGTGAAGGACGC 547
   |||

QY 168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
   |||
   |||
Db 548 GTGGAAGTGAAGCAGCAGAACCGTGTATCCGCTGTGTCGACAGAGCGTGGAGCTCCGG 607
   |||

QY 188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerPheIleSerSerGlyVal 207
   |||
   |||
Db 608 CCCGCAAGGTGGTGTGCGGCGAGCTGGGAGGTGTGGAGGCTGTACTGACTCAAAAAGC 667
   |||

QY 208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
   |||
```



668	TGCAACACAGCATCCCTATGGCGGAAMACCCATCTCGCTGAGCCAGATGTCATGCTCAGGA	721	
QY	227	AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis	246
Db	728	CGAAGACGAACCCCTTCAGGATGGCCCTCTCGGGCCTTSGGGGAAGAACAACCTGCAACCAT	787
QY	247	AsnGluAspValThrLeuThrCysTrpAspSerSerAspLeuGluLeuArgValGly	266
Db	788	GATGAAGACACGTGGGTGCAANTGTGAAGATCCCTTTGAC-----TTGAGACTAGTAGGA	841
QY	267	GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal	286
Db	842	GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCACAGGGCGTATGGGGCTCTGTC	901
QY	287	CysHisHisLysTrpAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGly	306
Db	902	TGTGATGACAACTGGGGAGAAAAGGAGGACCAAGGTGGTATGCAAGCAACTGGGCTGTGGG	961
QY	307	ThrAlaLeu-----HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal	324
Db	962	AAGTCCCTCTCTCCCTCTCTCAGACACCGGAATGCTATGGCCCTCTGGGGTTGGCCGCATC	1021
QY	325	TrpLeuAspGlyValSerCysSerGlyAsnGlnUerPheLeuTrpAspCysArgHisSer	344
Db	1022	TGCGTGGATAAATGTCCTGCTCAGGGGAGGAGCAGTCCCTGGACACTGCCACACAGA	1081
QY	345	GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer	362
Db	1082	TTTTGGGGGTTTTCCACGATGCACCCACAGGAAGATGTGGTGTCTATCTGCTCA	1135

-	PRIOR APPLICATION NUMBER:	US 60/146,222
:	PRIOR FILING DATE:	1999-07-28
:	PRIOR APPLICATION NUMBER:	PCT/US99/20594
:	PRIOR FILING DATE:	1999-09-08
:	PRIOR APPLICATION NUMBER:	PCT/US99/20944
:	PRIOR FILING DATE:	1999-09-13
:	PRIOR APPLICATION NUMBER:	PCT/US99/21090
:	PRIOR FILING DATE:	1999-09-15
:	PRIOR APPLICATION NUMBER:	PCT/US99/21547
:	PRIOR FILING DATE:	1999-09-15
:	PRIOR APPLICATION NUMBER:	PCT/US99/23089
:	PRIOR FILING DATE:	1999-10-05
:	PRIOR APPLICATION NUMBER:	PCT/US99/28214
:	PRIOR FILING DATE:	1999-11-29
:	PRIOR APPLICATION NUMBER:	PCT/US99/28313
:	PRIOR FILING DATE:	1999-11-30
:	PRIOR APPLICATION NUMBER:	PCT/US99/28564
:	PRIOR FILING DATE:	1999-12-02
:	PRIOR APPLICATION NUMBER:	PCT/US99/28565
:	PRIOR FILING DATE:	1999-12-02
:	PRIOR APPLICATION NUMBER:	PCT/US99/30095
:	PRIOR FILING DATE:	1999-12-16
:	PRIOR APPLICATION NUMBER:	PCT/US99/30911
:	PRIOR FILING DATE:	1999-12-20
:	PRIOR APPLICATION NUMBER:	PCT/US99/30999
:	PRIOR FILING DATE:	1999-12-20
:	PRIOR APPLICATION NUMBER:	PCT/US00/00219
:	PRIOR FILING DATE:	2000-01-05
:	NUMBER OF SEQ ID NOS:	423
:	SEQ ID NO 147	
:	LENGTH:	1686
:	TYPE:	DNA
:	ORGANISM:	Homo Sapien
:	US-09-904-011-147	
 Alignment Scores:		
Pred. No.:	2,64e-60	Length: 1686
Score:	691.50	Matches: 148
Percent Similarity:	52.51%	Conservative: 40
Best Local Similarity:	41.34%	Mismatches: 139
Query Match:	8.50%	Indels: 31
DB:	9	Gaps: 9
 US-09-759-1308-381 (1-1453) x US-09-904-011-147 (1-1686)		
Qy	21	LeuPheSerAlaValValThrCysIleLeuLeuAsnSerCys-----Phe 36
Db	107	CTATTTCC-----TTGATCCTTGCCATTTCACAGACCTGGATTC 148
Qy	37	LeuIleSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56
Db	149	CTAGCGTCTCCATCTGGA-----GTGCGCGCTGTGGGGGCCCTCCACCOC 193
Qy	57	CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspGly 76
Db	194	TGTCAAGGGCGGTGGAGGTGGAAACAGAACAGGCAGTGGGGCACCGTGTGTGATGACGCG 253
Qy	77	TrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
Db	254	TGGGACATTAAAGGACGTGGCTGTGTGTGCGGGAGCTGGGCTGTGGAGCTGCCACCGGA 313
Qy	97	Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
Db	314	ACCCCTAGTGGTATTTTGTATNGACCCACAGACAAAAGCAAANGTCTCTCATCCAA 373
Qy	114	AspValSerCystYrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
Db	374	TCAGTCAAGTCCACAGAACAGAAGATAFACATTGGCTCAGTGTGACGAGAAGAA----- 427
Qy	134	SerHisAnnCystYrHisGlyGluAspValGlyValAsnCys----- 147
Db	428	GTTTATCATGTTCCACATGATGAAGATGCTGGGGCATGCTGTGCAGAACCCAGAGACTCT 487

148	Qy	TyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArg	167
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
488	Db	TTCTCCCCAGTCCCAAGAGGGTGTCCAGCTGGCTCAGGCCCTGGCATTCGAAGGACGC	547
168	Qy	ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn	187
		:::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
548	Db	GTGGAGTGAACACCAAGCAGTGGTATACCGTGTGCCAGACAGCTGGAGCCTCCCG	607
188	Qy	ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerGlyVal	207
		:::: :::: :::: :::: :::: :::: :::: :::: ::::	
608	Db	GC CGCAAAAGTGTCTGCCGACGTGGATGTGGAGGCTGTACTGACTCAAAACGC	667
208	Qy	ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAlleLeuCysGlnGly	226
668	Db	TGCAACAAGCATGCCATATGGCGGAAAACCCATCTGGCTGAGCCAGATGTCATCTCAGGA	727
227	Qy	AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis	246
728	Db	CGAAGAAGCAACCTTCAGAGATTGCCCTTCTGGGCTTGGGGGAAGAACACCTGCAACCAT	787
247	Qy	AsnGluAspValThrLeuThrCysTrpAspSerSerAspLeuGluLeuArgLeuValGly	266
		:::       ::::	
788	Db	GATGAAGACACCTGGGTGCAATGTGAAGATCCCTTTGAC-----TTGACAGTAGTAGGA	841
267	Qy	GlyThrAsnArgCysMetClyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal	286
842	Db	GGAGAACACCTCTGCTGGGCGACTGGAGTGTGTGCACAAGGCGTATGGGCTCTGTC	901
287	Qy	CysHisHisLysTrpAsnAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGly	306
902	Db	TGTGATGACAACTGGGGAGAAAAGGAGGACCAAGTGTGTATGCAAGCAACTGGCTGTGG	961
307	Qy	ThrAlaLeu-----HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal	324
		:::	
962	Db	AAGTCCCTCTCCCTCCCTTCAGAGACCGGAATGCTATGCCCTGGGGTGTGCCGCATC	1021
325	Qy	TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer	344
1022	Db	TGGCTGGATATGTTCTGTTCTCAGGGGAGGAGCAGTCCCTGGAGCAGTCCGACACAGA	1081
345	Qy	GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer	362
1082	Db	TTTTTGGGGGTTTCAGCATGCAACCAAGAGATGGCTGTCTACTCTCTCA	1135

Search completed: May 12, 2003, 17:37:41  
Job time : 376 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 12, 2003, 06:37:56 ; Search time 446 Seconds  
(without alignments)  
7336.667 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWIDFGRCCHQN.....CEDASDTSLGLVPASEATK 1453

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US0759130/runat\_06052003\_123553\_23343/app\_query.fasta\_1.1607  
-DB=N\_Geneseq\_101002 -QFWT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US0759130.ecgn\_1.1.301 @runat\_06052003\_123553\_23343 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NRG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	8138	100.0	4359	22	Human TANGO 234 OR
2	8138	100.0	4628	22	Human TANGO 234 CD
3	6107.5	75.0	3670	22	DNA encoding human
4	4712	57.9	2800	22	Human full-length
5	3429	42.1	3716	22	Human polynucleoti
6	3429	42.1	3834	22	Human EST-derived
7	3429	42.1	3834	22	Human polynucleoti
8	3429	42.1	3834	22	Human polynucleoti
9	3396.5	41.7	3811	22	Human polynucleoti
10	3377	41.5	3707	24	Human DNA sequence
11	3082	37.9	4308	22	Bovine WCI ORF. B
12	2435	29.9	5802	19	Human SRCR protein
13	1763.5	21.7	2697	24	DNA encoding novel
14	1404	17.3	820	21	Human secreted exp
15	1183	14.5	690	22	Human protein enco
16	1142.5	14.0	4344	24	Rat sequence diffe
17	1142.5	14.0	4360	18	Rat von Ebner's gl
18	1081	13.3	608	20	EST clone Crl18.
19	1033	12.7	562	22	Human protein enco
20	972.5	12.0	1659	22	Human sbq14862SPER
21	921	11.3	1269	22	Human sbq14862SPER
22	830	10.2	2562	20	Human serine prote
23	830	10.2	3350	20	Human neurotrypsin
24	785.5	9.7	29598	19	Human Sc2 DNA. Ho
25	727	8.9	450	22	Mouse neurotrypsin
26	722	8.9	2376	20	Mouse serine prote
27	712	8.9	2614	20	Human SRCR protein
28	691.5	8.5	1686	20	Human SRCR protein
29	691.5	8.5	1686	20	Human SRCR protein
30	691.5	8.5	1686	24	Human SRCR protein
31	691.5	8.5	1686	24	Human SRCR protein
32	691.5	8.5	1686	24	Human SRCR protein
33	691.5	8.5	1686	24	Human SRCR protein
34	691.5	8.5	1686	24	Human SRCR protein
35	691.5	8.5	1686	24	Human SRCR protein
36	691.5	8.5	1686	24	Human SRCR protein
37	684	8.4	527	22	Human SRCR protein
38	647.5	8.0	3767	21	Human SRCR protein
39	639.5	7.9	3532	24	Human SRCR protein
40	638	7.8	2420	22	Human SRCR protein
41	635.5	7.8	5059	22	Human SRCR protein
42	634.5	7.8	2262	21	Human SRCR protein
43	634.5	7.8	2262	21	Human SRCR protein
44	634.5	7.8	2785	24	Human SRCR protein
45	634.5	7.8	2920	21	Human SRCR protein

# ALIGNMENTS

RESULT 1

AAF45124

ID AAF45124 standard; cDNA; 4359 BP.

AC AAF45124;

DT 30-MAR-2001 (first entry)

DE Human TANGO 234 ORF.

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;  
XX central nervous system; focal brain disorder; bipolar affective disorder;  
XX global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;  
XX senile dementia; Huntington's disease; amyotrophic lateral sclerosis;  
XX Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep.  
XX neuropsychiatric; psychoactive substance use; anxiety; ss.  
OS Homo sapiens.

XX WO200077239-A2.  
PN  
XX  
PN  
PD 21-DEC-2000.  
XX  
PF 24-MAY-2000; 2000WO-US14858.  
XX  
PR 14-JUN-1999; 99US-0333159.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;  
XX WPI; 2001-032313/04..  
XX P-PSDB; AAB66037.  
DR  
XX  
PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for  
PT screening assays and diagnostic assays and for the treatment of  
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's  
PT disease -  
XX  
PS Claim 1; Fig 2; 359pp; English.  
XX  
CC The present invention relates to TANGO or INTERCEPT proteins and coding  
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,  
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding  
CC sequences are useful for the treatment of neurological disorders such as  
CC central nervous system (CNS) disorders, CNS-related disorders, focal  
CC brain disorders, global-diffuse cerebral disorders and other  
CC neurological and cerebrovascular disorders. The CNS disorders include  
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic  
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,  
CC autonomic function disorders such as hypertension and sleep disorders,  
CC neuropsychiatric disorders, psychoactive substance use disorders,  
CC anxiety, and bipolar affective disorder.  
XX  
SQ Sequence 4359 BP; 1060 A; 881 C; 1303 G; 1115 T; 0 other;

Alignment Scores:  
Pred. No.: 0 Length: 4359  
Score: 8138.00 Matches: 1453  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 22 Gaps: 0

US-09-759-130B-381 (1-1453) x AAF45124 (1-4359)

QY 1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHisGlnAsn 20  
DB 1 ATGATGCTGCCTCAAAACCTCGTGCATATTGATTTTGGAGATGCTGCTCATCAGAAC 60  
QY 21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40  
DB 61 CTTTTCCTGCTGTGTAACCTGCATCCCTGCTCCTCAATTCCTGCTTCATCAGCAGT 120  
QY 41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60  
DB 121 TTTAATGGACAGATTGGAGTGGAGCTGGTGGTCAATGAGAGCGTCCCTGCTGGACA 180  
QY 61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80  
DB 181 GTGGAGGTGAAATTCACGAGACAGTGGGGAGCTGTGTGTGATGATGGGTGGACACTACT 240  
QY 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100  
DB 241 GCCTCAACTGCTGTGTCACAAACAGCTGGATGTCCTTCCTTCGCCAAGTTTCGTTTT 300  
QY 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysThrGlyAsn 120  
DB 301 GGACAGCGGTAGTACATGACATGAAAAATTTGGCTGTGATGATGTTTCCCTGTTATGAAAT 360  
QY 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysThrHisGly 140

DB 361 GAGTCAGCTCTCTGGGAATGTCAACACCGGAATGGGAAGCCATAACTGTATCATGSA 420  
QY 141 GluAspValGlyValAsnCysTrpGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160  
DB 421 GAAGATGTTGGTGTGAACCTGTATGTTGAAGCAATCTCTGGGTTTGGGCTAGTGATGA 480  
QY 161 AsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCys 180  
DB 481 AACAACTCCTGTTTCAGGAGAGTGGAGTGAAATTCAGAAAGGTTGGGGACTATATCT 540  
QY 181 AspAspGlyTrpAsnLeuAsnThrAlaValValCysArgGlnLeuGlyCysProSer 200  
DB 541 GATGATGGTGGAACTTGAATCTCTCCGCTGGTGTGAGCAACTAGGATGTCCATCT 600  
QY 201 SerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAsp 220  
DB 601 TCTTTATTTCTCTCGGAGTGTAAATAGCCCTGCTGTATTGCGGCCCATTTGGCTGGAT 660  
QY 221 AspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGly 240  
DB 661 GACATTTTATGCCAGGGGAATGAGTTGGCACTCTGGAATTCAGACATCGTGGATGGGA 720  
QY 241 AsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTrpAspSerSerAspLeu 260  
DB 721 AATCATGACTGCAGTCACAAATGAGGATGTCACTTAACCTGTTATCATAGTAGTATCT 780  
QY 261 GluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGln 280  
DB 781 GAACTAAGCTTGTAGTGGAACTAACCCCTGTATGGGAGAGTAGAGCTGAAATCCAA 840  
QY 281 GlyArgTrpGlyThrValCysHisIleLysTrpAsnAsnAlaAlaAspValValCys 300  
DB 841 GCAAGGTGGGGACCGTATGCCACCATAGTGGCAATGCTGCACTGATGCTGTATGC 900  
QY 301 LysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGly 320  
DB 901 AAGCAGTTGGGATGTGGAACCCACTTCACTTCGCTGGCTTGCCTCATTTTGCAGTCAGG 960  
QY 321 SerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAsp 340  
DB 961 TCTGATGTTGATGCTGTGATGCTCTCTCGGTAATGTAATCTTTCTTTGGGAC 1020  
QY 341 CysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIle 360  
DB 1021 TGCAGACATTCGGAACCGTCAATTTTGACTGCTCTCATCAAAACGATGCTGTGTATC 1080  
QY 361 CysSerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGly 380  
DB 1081 TGCTCAGATGGAGCAGATTGGAACTGCCAGCTAGCAGATGGAGATGAACATTTGTCAGG 1140  
QY 381 ArgValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsn 400  
DB 1141 AGAGTAGAGTGCAGATTTCATGACAGTGGTGAACAATATGTGACCAGAACTGGAAGAAT 1200  
QY 401 GluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArg 420  
DB 1201 GAACAAGCCCTGTGTTGTAAGCAGCTAGGATGTCCCTTCAGCGCTCTTTGGCAGCTGT 1260  
QY 421 ArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGly 440  
DB 1261 CGTGCTAAACCTAGTAATGAAGTAGAGCATTTGGATTAACAGCATATCTTGCACCTGG 1320  
QY 441 AsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArgThrCysPheArg 460  
DB 1321 AATGACTCAGCTCTCTGGGAGCTGCATATGATGGAAGAAAGCAAGCAATGCTTCCGA 1380  
QY 461 ArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGly 480  
DB 1381 AGATCAGATGCTGGAGTAATTTCTTGATAAGGCGAGATCTGGACCTAAGGCTGTGCGG 1440  
QY 481 AlaHisSerProCysTrpGlyArgLeuGluValLysTrpGlnGlyGluTrpGlyThrVal 500



```
QY 1221 GlnCysLeuSerAlaProTrpGluArgIleSerSerProAlaGluThrTrpIle 1240
DB 3661 CAGTGCCTCTCTGCCCATGGAGGAGAGAAATCTCAGCCAGAGAGACCTGGATC 3720
QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyValGlu 1260
DB 3721 ACATGTGAAGATAGAATAGAGTGGCTGGAGGAGACCGAGTCTCTGGAGAGTGGAG 3780
QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAla 1280
DB 3781 ATCTGGCAGCGAGCTCTCTGGGACAGTGTGTGATGACTCTGGGACCTGCCGAGGG 3840
QY 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300
DB 3841 GAAAGTGGTGTGACAGAGTGGCTGTGGCTCTGTCTGTCTGTCTGTCTGTCTGTCTG 3900
QY 1301 PheGlyGlnGlyThrClyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSer 1320
DB 3901 TTTGGCCAGGGAACCTGAACCATCTGGTGGATGACATCGGTGCAAGAGAAATGAGTCA 3960
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340
DB 3961 TTTCTATGGGACTGTACACCCAAACCTGGGACAGAGTGTGGACACAGGAAGAT 4020
QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360
DB 4021 GTGGCGTGTGAGTGTCTGTGACAGTGTCTGAAATCACTGAATGCCTCTCAGTCAATTA 4080
QY 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380
DB 4081 GCACATATTTATCCAGTATCTTTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 4140
QY 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArg 1400
DB 4141 ACCTGTGTCGAGTTCAGAAACAAACAACTCTGCCCTCAGAGTTTCAACCAAGAGGAGG 4200
QY 1401 GlySerLeuGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420
DB 4201 GGTTCCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCTCAAGAGAGGACCCA 4260
QY 1421 HisGlyThrArgThrSerAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440
DB 4261 CATGGACAAAGAACCTCAGATCACACCCCAACCATGGTTGTGAAGATGCTAGCGACACA 4320
QY 1441 SerLeuLeuGlyValLeuProLaserGluAlaThrLys 1453
DB 4321 TCGCTGTGGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4359
RESULT 2
AAF45123
ID AAF45123 standard; cDNA; 4628 BP.
XX
AC AAF45123;
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 234 cDNA.
XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS Homo sapiens.
XX
PN WO200077239-A2.
XX
PD 21-DEC-2000.
XX
PF 24-MAY-2000; 2000WO-US14858.
XX
```

```
PR 14-JUN-1999; 99US-0333159.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX
WPI; 2001-032313/04.
DR P-PSDB; AAB66037.
XX
TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
screening assays and diagnostic assays and for the treatment of
neurological diseases such as Alzheimer's, Parkinson's and Huntington's
disease -
XX
Claim 1; Fig 2; 359pp; English.
XX
The present invention relates to TANGO or INTERCEPT proteins and coding
sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
sequences are useful for the treatment of neurological disorders such as
central nervous system (CNS) disorders, CNS-related disorders, focal
brain disorders, global-diffuse cerebral disorders and other
neurological and cerebrovascular disorders. The CNS disorders include
Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
autonomic function disorders such as hypertension and sleep disorders,
neuropsychiatric disorders, psychoactive substance use disorders,
anxiety, and bipolar affective disorder.
XX
SQ Sequence 4628 BP; 1174 A; 925 C; 1352 G; 1177 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 4628
Score: 8138.00 Matches: 1453
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-759-130B-381 (1-1453) x AAF45123 (1-4628)
QY 1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHisGlnAsn 20
DB 28 ATGATGTCGCCCAAACTCGGGCATATGATTTGGAGATGCTGCTGTCATCAGAAC 87
QY 21 LeuPheSerAlaValValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSer 40
DB 88 CTTTCTCTGCTGTGTAACTTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 147
QY 41 PheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60
DB 148 TTTAATGGAAACAGATTGGAGTTGAGGCTGGTCAATGGAGACGGTCCCTCTGGGACA 207
QY 61 ValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThrThr 80
DB 208 GTGGAGTGAATTCACGAGCAGTGGGGGACTGTGTGTGTGTGTGTGTGTGTGTGTGT 267
QY 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArgPhe 100
DB 268 GCCTCAACTGCTCGTGGCAACAGCTTGGATGTCCATTTCTTTCGCCATGTTTCGTTTT 327
QY 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTrpGlyAsn 120
DB 328 GGACAGCGCGTACTAGACATGGAATAATTTGGCTTGTGTGTGTGTGTGTGTGTGTGT 387
QY 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTrpHisGly 140
DB 388 GAGTCAGCTCTCTGGGAATGTCACACCCGGATGGGAGGCCATACCTGTTATCATGGA 447
QY 141 GluAspValGlyValAsnCysTrpGlyGluAlaAsnLeuGlyLeuArgLeuValAspGly 160
DB 448 GAAGATGTTGGTGAACCTGTTATGTTGAGCAATCTGGGTTGAGGCTAGTGGATGGA 507
```



Db 2668 GAAGACACTTGTATCCACAGCAGAGAAGTTGGAGTTGTCTGTGCCGATATACAGATGTC 2727  
QY 901 ArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHis 920  
Db 2728 CGACTTGTGANTGGCAATCCAGAGTGTGACGGGCAAGTGGAGATCAACGTGCTTGGACAC 2787  
QY 921 TrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaAargValLeuCysArgGln 940  
Db 2788 TGGGGCTCACTGTGTGACACCCACTGGGACCCAGAGAGTGGCCGTGTTCTATGACAGACAG 2847  
QY 941 LeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrlleGlyGluArgSerVal 960  
Db 2848 CTCAGCTGTGGGACTGCTCTCAACACACAGGAGAAATATATTGGAGAAAGAGTGT 2907  
QY 961 ArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGln 980  
Db 2908 CGTGTGTGGGACACAGGTTTCATTGCTTAGGGAATGAGTCACTCTGGATTAACGTCAA 2967  
QY 981 MetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThr 1000  
Db 2968 ATGACAGTTCTTGGAGCACTCCCTGTATCCATGGAAATACTCTCTGTGATCTGCACA 3027  
QY 1001 GlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrlleuSer 1020  
Db 3028 GGAAGCCTCAACAGGCACTGTTTCATGCCCTGCAAAATGTATCTGACCCATATTGTC 3087  
QY 1021 AlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuValAsp 1040  
Db 3088 GCAGTTCCAGAGGGCAGTCTTTGATCTCTTAGAGGACAAACGGCTCCGCCCTAGTGGAT 3147  
QY 1041 GlyAspSerArgCysAlaGlyArgValGluIleTyrlleHisAspGlyPheThrGlyThrIle 1060  
Db 3148 GGGGACAGCGCTGTGCCGGGAGAGTAGAGATCTATCACAGCGCTCTTGGGGCACCATC 3207  
QY 1061 CysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGly 1080  
Db 3208 TGTGATGACGGCTGGGACCTGACGATGCCAGTGGTGTGTCAAAAGCTGGGCTGTGGA 3267  
QY 1081 ValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlyserGlyProIleTrpLeu 1100  
Db 3268 GTGGCCTTCAATGCCACGCTCTGTCTACTTTGGGGAGGGGTTCAGGGCCCATCTGGCTG 3327  
QY 1101 AspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrp 1120  
Db 3328 GATGACCTGAATGCAACAGGAAGGAGTCCCATCTGTGCAGTGCCTTCCCGCGGCTGG 3387  
QY 1121 GlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAla 1140  
Db 3388 GGGCACACGACTGCAGGCACAGAGGAGGAGCGGGGTCTATCTGCTCAGAAATTCACAGCC 3447  
QY 1141 LeuArgLeuTyrlleSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrl 1160  
Db 3448 TTGAGGCTCTACAGTGAATGAAACAGAGAGTGTGCTGGGAGATTTGGAAGTCTTCTAT 3507  
QY 1161 AsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleVal 1180  
Db 3508 AACGGACCTTGGGGACGCTGGCAGGAGGAGACATCACCACAGCCATAGCAGCATTTGG 3567  
QY 1181 CysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThr 1200  
Db 3568 TGCAGGCAGCTGGGCTGTGGGAGATGAGTGTGAGCTCGCCCTTATCTAAGACA 3627  
QY 1201 GlySerGlyPheMetTrpValAspAlleGlnCysProLysThrHisIleSerIleTrp 1220  
Db 3628 GGCTCTGGTTTCATGTGGGTGGATGATTCAGTGTCTTAAACGCATATCTCCATATGG 3687  
QY 1221 GlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIle 1240  
Db 3688 CAGTGGCTGTCTGCCCATGGGAGCGAAGAAATCTCAGGCCCGCAGCAAGAGACCTGGATC 3747  
QY 1241 ThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260  
Db 3748 ACATGTGAAGATAGANTAGAGTGCCTGGGAGGAGACCGAGTGTCTGGGAGATGGAG 3807

QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280  
Db 3808 ATCTGGCAGCAGGCTCTCTGGGGCACAGTGTGTGATGACTCTCTGGACCTGCCGAGGCG 3867  
QY 1281 GluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300  
Db 3868 GAAGTGGTGTGTGACAGCTGGGCTGTGGCTCTGTCTGTCTGCTGAGGAGACGCTTCG 3927  
QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320  
Db 3928 TTTGGCCAGGGAACCTGGAAACCATCTGGTTGGATGACATCGGTGCAAGAGAAATGAGTCA 3987  
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340  
Db 3988 TTTCTATGGGACTGTACGCCAAACCTCTGGGACAGAGTACTGTGGACACAAGGAAGAT 4047  
QY 1341 AlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeu 1360  
Db 4048 GCTGGCGTGAAGTGTCTGGACAGTCGCTGAAATCACTGAATGCCCTCTCAGGTCAATTA 4107  
QY 1361 AlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeu 1380  
Db 4108 GCACCTATTATCCAGTATCTTTGGGCTCCTCTCTCTGTTCTGTTATTCTATTCTC 4167  
QY 1381 ThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArg 1400  
Db 4168 ACGTGTGTCGAGATTCAGAAACAAACATCTGCCCTCAGAGTTTCAACACAGAGAGG 4227  
QY 1401 GlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspPro 1420  
Db 4228 GGTCTCTCGAGAGAAATTTATTCATGAGATGGAGACCTGCTCAAGAGAGGAGGCCA 4287  
QY 1421 HisGlyThrArgThrSerAspThrProAsnHisGlyCysGluAspAlaSerAspThr 1440  
Db 4288 CATGGACAAGAACCTCAGATCAGACACCCCAACCACTGTTGTGAAGATCTAGCGACACA 4347  
QY 1441 SerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453  
Db 4348 TCGCTGTGGGAGTCTCTCTGCTGCTGAGGCCACAAAA 4386  
RESULT 3  
AS01217  
ID AAS01217 standard; cDNA; 3670 BP.  
XX AAS01217;  
DT 04-JUL-2001 (first entry)  
DE DNA encoding human secreted protein, POLY8.  
XX Human secreted protein; therapeutic; diagnostic; human; cancer; ss.  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 113..3475  
FT /tag= a  
FT /product= "Human secreted protein POLY8"  
XX  
XX WO200119856-A2.  
XX  
XX 22-MAR-2001.  
XX  
XX 13-SEP-2000; 2000WO-US25106.  
XX  
XX 13-SEP-1999; 99US-0153629.  
PR 16-SEP-1999; 99US-0154520.  
PR 20-SEP-1999; 99US-0154762.  
PR 13-OCT-1999; 99US-0159231.  
XX 12-SEP-2000; 2000US-0659634.  
XX  
XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Fernandes E, Herrmann JL, Liu X, Yang M, Boidog FL;  
 XX WPI; 2001-244781/25.  
 DR P-PSDB; AAU00396.  
 XX  
 XX  
 PT New POLYX polypeptide useful for treating or preventing a POLYX  
 PT associated disorder, e.g. cancer -  
 PS  
 PS Claim 9; Page 25-29; 152pp; English.  
 XX  
 XX The sequence represents the coding sequence of human secreted protein,  
 CC POLY8. POLYX nucleic acids, polypeptides and antibodies to POLYX can be  
 CC used for treating or preventing a POLYX associated disorder in a subject,  
 CC preferably a human. These can be used in the manufacture of a medicament  
 CC for treating a syndrome associated with a human disease selected from a  
 CC POLYX-associated disorder, where the therapeutic is a POLYX polypeptide,  
 CC a POLYX nucleotide or a POLYX antibody. They may also be used to screen  
 CC for a modulator of activity, or latency, or predisposition to a POLYX  
 CC associated disorder, e.g. cancer.  
 XX  
 SQ Sequence 3670 BP; 947 A; 752 C; 1073 G; 898 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 3670  
 Score: 6107.50 Matches: 1103  
 Percent Similarity: 95.33% Conservative: 0  
 Best Local Similarity: 95.33% Mismatches: 3  
 Query Match: 75.05% Indels: 51  
 DB: 22 Gaps: 1  
 US-09-759-130b-381 (1-1453) x AAS01217 (1-3670)  
 QY 348 AsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer----- 362  
 DB 2 AATTTTCACTGCTCTCATCAAAAGATGCTGCTGTGATCTGCTCAGACCAGTTGCTATTC 61  
 QY 362 ----- 362  
 DB 62 AAACCTCTGGAACCTCTGGAACCTATTTGCCCTCTGGAAGAGTAAGGAGGCCCATGAAGAAT 121  
 QY 362 ----- 362  
 DB 122 GATGCAGATGTGAAATAAGTTGGAAACCGTCCCAACAGTGAAGAACAGATGTGTATGG 181  
 QY 363 -----AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsn 376  
 DB 182 CCAACCCACAGGAAATGGATGGAGCAGATTGGAACTGCGACTAGCAGATGGAAGTAAC 241  
 QY 377 AsnCysSerGlyArgValGluValArgIleHisGluGlnTrpThrIleCysAspGln 396  
 DB 242 AATGTTCAGGAGAGTAGAGTGAGAATTCATGAACAGTGGTGGCAATATGTGACCAG 301  
 QY 397 AsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerVal 416  
 DB 302 AACTGGAAGATGAACAGCCCTTGTGGTTGTGAAGCAGCTAGGATGCTCCGCTTCAGCGTC 361  
 QY 417 PheGlySerArgArgAlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIle 436  
 DB 362 TTGGCCAGCTGCTGCTGCTAAACCTAGTAATGAAGCTAGACACATTTGGATTAACAGCATA 421  
 QY 437 SerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTrpAspGlyLysAlaLysArg 456  
 DB 422 TCTTGCACTGGGAATGATGAGTCAGCTCTCTGGGACTGCACATATGATGAAAGCAACGGA 481  
 QY 457 ThrCysPheArgSerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeu 476  
 DB 482 ACATGCTTCGAGATCAGATGCTGGAGTAATTTGTTCTGTATAGGAGCAGATCTGACCTA 541  
 QY 477 ArgLeuValGlyAlaHisSerProCysTyrClyArgLeuGluValLysTyrGlnGlyGlu 496  
 DB 542 AGGCTTCTGGGGCTCATAGCCCTGTTATGGGAGATTGGAGGTGAATACCAAGAGAGAG 601

QY 497 TrpGlyThrValCysHisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGln 516  
 DB 602 TGGGGAGCTGTGTCTCATGACAGATGAGCAAGGAATGACAGCTGTGTGTGTAACAA 661  
 QY 517 LeuGlyCysGlyLysProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGly 536  
 DB 662 TTGGGATGTGGAAGCCATGATGATGTGTGTGTATGACCTATTATTAAGAAGACATCAGGA 721  
 QY 537 ProIleTrpLeuAspAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGlu 556  
 DB 722 CCTATTGGCTGGATGACGTTCTTGCAATGGAATGAGTCAATATCTGGGAGTGA 781  
 QY 557 HisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCysSer 576  
 DB 782 CACAGTGGATGGGAAAGCATAATTGTTGTACACAGAGAGGATGTGATTTAACTGTCTCA 841  
 QY 577 GlyAspAlaThrTrpGlyLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeu 596  
 DB 842 GGTGATCAACATGGGCGCTGAGGCTGTGGCGGCGACACCCGCTGCTCGGGAAGACTG 901  
 QY 597 GluValTyrPheGlnGlyArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAla 616  
 DB 902 GAGGTGTACTTCAAGGACGCTGGGCGACAGTGTGTGATGACGCTGGAACGGTAAGCT 961  
 QY 617 AlaAlaValValCysSerGlnLeuAspCysProSerIleIleGlyMetGlyLeuGly 636  
 DB 962 GCAGCTGTGGTGTGTAGCCAGCTGGACTGCCCATCTTCTATCATTTGGCATGGGTGGGA 1021  
 QY 637 AsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGlu 656  
 DB 1022 AACGCTTCTACAGGATATGGAAATTTGGCTCGATGATGTTCTCTGTGATGGAGATGAG 1081  
 QY 657 SerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGlu 676  
 DB 1082 TCAGATCTCTGCTCATGCGAAGACAGTGGGTGGGAAATAATGACTGCGACGTACAGTGA 1141  
 QY 677 AspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySer 696  
 DB 1142 CATCTTGGAGTGATCTGTTCTGATGTCATGCGATATGAGGCTGAGGCTTGTGGTGAAGC 1201  
 QY 697 SerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAla 716  
 DB 1202 AGCAGGTGTGCTGGAAGATTGAGTGAATGTCAGGCTGCCGTGGGAATTTCTGTGTCT 1261  
 QY 717 AsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAla 736  
 DB 1262 AATGCTGGGGAATGAACATTTGCTGAAGTCTTTGCAGGCAACTTGAATGTGGGTCTGCA 1321  
 QY 737 IleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsn 756  
 DB 1322 ATCAGGCTCTCCAGAGAGCCCTCAITTCACAGAAAGAACATTACACATCTTAATGTGGAAT 1381  
 QY 757 SerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGln 776  
 DB 1382 TCTGGCTGCACTGGAGGGAGCCCTCTCTGGGATTGTATACGATGGGATGGGAACAG 1441  
 QY 777 ThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArg 796  
 DB 1442 ACTCGCTGTCATTTAATATATGAAGCAAGTTTGTATGCTGAGCCACAGCCAGCCAG 1501  
 QY 797 LeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrp 816  
 DB 1502 CTGTTGGAGCTGATATGCCCTGCTCTGAGCGTGTGAAGTGAACATGCGACACATG 1561  
 QY 817 ArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeu 836  
 DB 1562 CGCTCTCTGTGTGATTTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1621  
 QY 837 AsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLys 856  
 DB 1622 AACTGTGGAGATGCCATATCTTCTTCTGTTGGAGATCATCTTTGGAAAGGAATGGTCTA 1681  
 QY 857 ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIle 876

Db 1682 ACTTGGCCGAAAGTTCAGTGTGAAGGAGTGAACCTTCACCTTGCATTTATGCCCCATT 1741  
QY 877 ValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArg 896  
Db 1742 GTTCAACATCCGAGACACTGTATCCACACAGAGAAGTTGGAGTGTCTGTCCCCGA 1801  
QY 897 TyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsn 916  
Db 1802 TATACAGATGTCGACTTGTGAATGGCAATCCCAAGTGCAGCGGCAAGTGGAGATCAAC 1861  
QY 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936  
Db 1862 GTGCTGGACATGGGGCTCACTGTGTGACACCCACTGGGACCCAGAAAGATCCCGTGT 1921  
QY 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTrpIleGly 956  
Db 1922 CTATGCAGACACTCAGCTGTGGACTGCTCTCAACACAGAGGAAATATATTGA 1981  
QY 957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976  
Db 1982 GAAAGAAGTGTTCGTGTGTGGGGACACAGGTTTCATTTGCTTAGGGAATGAGTCACTTCTG 2041  
QY 977 AspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSer 996  
Db 2042 GATACTGTCAATGACACTTCTGGAGCACCTTCCCTGTATCCATGGAAATACTGTCTCT 2101  
QY 997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016  
Db 2102 GTGATCTGCACAGGAAGCTGACCAGCCACTGTTTCCATGCTCGCAATGTATCTGAC 2161  
QY 1017 ProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036  
Db 2162 CCATATTTCTGTGAGTCCAGAGGCGAGTGTGTTGATCTGTAGAGGACAAAGGGCTC 2221  
QY 1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleThrHisAspGlyPhe 1056  
Db 2222 CGCCTAGTGGTGGGACAGCGCTGTGCCGGAGAGTAGAGATCTATCACCAGCGCTTC 2281  
QY 1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076  
Db 2282 TGGGGCACCACATGTGATCAGCGCTGGGACCTGAGCGATGCCCGTGGTGTGTCAAAG 2341  
QY 1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096  
Db 2342 CTGGGCTGTGAGTGGCCCTTCAATGCCAGGCTCTGCTCACTTTGGGGAGGGGTCAAGG 2401  
QY 1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116  
Db 2402 CCATCTGGCTGGTGAACCTGAACCTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCCT 2461  
QY 1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136  
Db 2462 TCCCGCGCTGGGGCAGCACCACCTGCAGCCACAGGAGGAGCGAGGGTCACTCTCTCA 2521  
QY 1137 GluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156  
Db 2522 GAATTCACAGCCTTGAAGGCTTACAGTGAACCTGAAACAGGAGGAGTGTCTGGGAGATTG 2581  
QY 1157 GluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176  
Db 2582 GAAGTCTTCTATAAGCGGACCTGGGGCAGCGTGGCAGGAGGAAACATCACACAGCCATPA 2641  
QY 1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196  
Db 2642 GCAGCATTTGTGCAGGCAGCTGGGCTGTGGGAGATGGAGTGTCAACCTCGCCCT 2701  
QY 1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216  
Db 2702 TTATCTAAGACAGGCTCTGGTTTCTATGTGGTGGATGACATTCAGTGTCTTAAACGCAT 2761  
QY 1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236

Db 2762 ATCTCCATATGCGAGTGCCTGTCTGCCCCCATGGGACGGAAGAAATCTCCAGCCAGCAGAA 2821  
QY 1237 GluThrTrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSer 1256  
Db 2822 GAGACTGGATCACATGTCAAGATAGAATAAGAGTGCCTGGAGGAGACACCGAGTCTCT 2881  
QY 1257 GlyArgValGluIleThrPheHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAsp 1276  
Db 2882 GGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGGGCACAGTGTGTATGACTCTCTGGGAC 2941  
QY 1277 LeuAlaGluAlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeu 1296  
Db 2942 CTGGCCGAGCGGAGTGTGTGTGACAGCTGGGCTGTGCTCTCTGGCTGGCCCTG 3001  
QY 1297 ArgAspAlaSerPheClyGlnGlyThrGlyThrIleTrpLeuAspMetArgCysLys 1316  
Db 3002 AGGGACGCTTCGTTTGGCCAGGAACTGGAACCATCTGTTGGATGACATGCGGTGCAAA 3061  
QY 1317 GlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGly 1336  
Db 3062 GGAATATGAGTCAATTTTATGGGACTGTACAGCAACCTGGGGACAGAGTCACTGTGA 3121  
QY 1337 HisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSer 1356  
Db 3122 CACAAGGAAGATGCTGGCGTGAGGTCTCTGGACAGTCTGAAATCACTGAATGCTCC 3181  
QY 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376  
Db 3182 TCAGGTCAATTTAGCACTTATTTATCCAGTATCTTTGGGCTCTCTCTCTGTTCTGTT 3241  
QY 1377 IleLeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSer 1396  
Db 3242 ATTCTATTTCTACGCTGGTGGCGAGTTCAGAAACAAAAACATCTGCCCTTCAGAGTTCA 3301  
QY 1397 ThrArgArgArgGlySerLeuGluAsnLeuPheHisGluMetGluThrCysLeuLys 1416  
Db 3302 ACCAGAGAGGGGTCTCTCGAGGAGAAATTTATCCATGAGATGGAGACCTGCCCTCAAG 3361  
QY 1417 ArgGluAspProHisGlyThrArgThrSerAspAspThrProAsnHisGlyCysGluAsp 1436  
Db 3362 AGAGAGGACCCACATGGGACAAAGAACCTCAGATGACACCCCAACCATGTTGTGAAGAT 3421  
QY 1437 AlaSerAspThrSerLeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453  
Db 3422 GTAGCGACACATCGCTGTTGGAGTTCTTCTGCTGCTCTGAAGCCACAAA 3472  
RESULT 4  
AAS44609  
ID AAS44609 standard; DNA; 2800 BP.  
XX  
AC AAS44609;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human full-length polynucleotide sequence #34.  
XX  
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;  
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;  
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;  
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
KW cytostatic; antirheumatic; antiallergic; vulnary; antiinflammatory;  
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;  
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;  
KW immunostimulant; analgesic; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200164834-A2.  
XX  
PD 07-SEP-2001.  
XX

PF 26-FEB-2001: 2001WO-US04926.  
 XX 28-FEB-2000: 2000US-0515126.  
 PR 18-MAY-2000: 2000US-0577409.  
 PR 17-JUN-2000: 2000US-0597707.  
 PR 14-JUL-2000: 2000US-0616807.  
 PR 19-SEP-2000: 2000US-0664641.  
 XX (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
 PI Drmanac R;  
 XX  
 DR WPI: 2001-589862/66.  
 DR P-PSDB: AAU27709.  
 XX  
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
 prepared from various human tissues, for diagnosis, treatment of  
 PT cancer, neurological, inflammatory disorders and for use in arrays for  
 PT detection  
 XX  
 PS Claim 1: SEQ ID NO 34; 153pp; English.  
 XX  
 CC Sequences AAS44576-AAS44919 represent full-length polynucleotides and  
 CC contig polynucleotides encoding polypeptides of the invention. The DNA  
 CC and protein sequences are useful for the treatment, diagnosis and  
 CC prevention of various types of disorder in a mammalian subject such as a  
 CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers  
 CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such  
 CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,  
 CC cell proliferation, cell differentiation, stem cell growth factor,  
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells  
 CC in culture to give rise to neuroepithelial cells that can be used to  
 CC augment or replace cells damaged by illness, accidental damage or genetic  
 CC disorders. The sequences may also be used for regeneration of bone,  
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.  
 CC Note: Some sequences for this patent did not form part of the printed  
 CC specification, but were obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 2800 BP; 697 A; 607 C; 813 G; 683 T; 0 other;

Alignment Scores:  
 Pred. No.: 3.57e-293 Length: 2800  
 Score: 4712.00 Matches: 850  
 Percent Similarity: 99.77% Conservatives: 0  
 Best Local Similarity: 99.77% Mismatches: 2  
 Query Match: 57.90% Indels: 0  
 DB: 22 Gaps: 0

US-09-759-130B-381 (1-1453) x AAS44609 (1-2800)

Qy 602 GlyArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCys 621  
 Db 37 GGACGGTGGGCACAGCGTGTGATCACGGCTGGAACAGTAAGCTGCAGCTGTGGTGTGT 96  
 Qy 622 SerGlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGly 641  
 Db 97 AGCCAGCTGGAGCTGCCCATCTCTATATTGGCATGGCTCTGGGAAAGCTCTACAGGA 156  
 Qy 642 TyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSerAspLeuTrpSer 661  
 Db 157 TATGAAAAATTTGGCTCGATGATGTTCTCTGTCGATGAGATGATGATGATGATGATG 216  
 Qy 662 CysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValIcylValile 681  
 Db 1297 GTTCCAGAGGGCAGTGTCTTGTATCTGCTTAGAGGACAAACGGCTCCGCTAGTGTGGTGG 1356

Db 217 TGCAGGAACACGTGGGTGGGAAATAATGACTGCAGTCACAGTGAAGATGTTGGAGTGATC 276  
 Qy 682 CysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGly 701  
 Db 277 TGTTCCTGATGATCGGATATGAGCTGAGGCTTGTGGGTGGAAGCAGCAGGTGCTCGGA 336  
 Qy 702 LysValGluValAsnValGlnGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMet 721  
 Db 337 AAAGTTGAGGTGAATGTCAGGGTGGCGTGGGAATTCGTGTCTAATGGCTGGGAATG 396  
 Qy 722 AsnIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArg 741  
 Db 397 AACATTGCTGAAGTTGTTTGCAGGCAACTTGAATGTGGTCTGCAATCAGGCTCTCCAGA 456  
 Qy 742 GluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGly 761  
 Db 457 GAGCCTCATTTTCACAGAAAGAACATTTACATCTTAATGTGCAATTCGTGGCTGCACTGA 516  
 Qy 762 GlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeu 781  
 Db 517 GGGGAAGCCCTCTCTGGGATTGTATACGATGGGAGTGGAAACAGACTGCGTGTCAATTA 576  
 Qy 782 AsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValIcylAlaAsp 801  
 Db 577 AATATGGAAGCAAGCTTTGATCTGCTACGCCACAGCCAGCCAGGCTGGTTGGAGCTGAT 636  
 Qy 802 MetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAsp 821  
 Db 637 ATGCCCTGCTCTGGACGTGTGAAGTGAACATGCAGACACATGCGCGTCTGTCTGTGAT 696  
 Qy 822 SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla 841  
 Db 697 TCTGATTCTCTCTTCATGCTGCCAATGCTGTGTCAGAGAATTAACCTGTGGAGATGCC 756  
 Qy 842 IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLys 861  
 Db 757 ATATCTCTTCTGTGGGAGATCACTTTGGAAGGAAATGTTCTAACTTGGCCGCAAAAG 816  
 Qy 862 PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu 881  
 Db 817 TTCAGTGTGAAGGAGTGAACCTCACCTTGCATTATGCCCCATGTTTCAACATCCGGA 876  
 Qy 882 AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArg 901  
 Db 877 GACACTTGTATCCACAGCAGAGAAGTTGGAGTTGCTCTTCCCGATATACAGATGTCGA 936  
 Qy 902 LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp 921  
 Db 937 CTGTGTAATGGCAATCCAGTGTGACGGCAAGTGGAGATCAACGTGCTTGGACACTGG 996  
 Qy 922 GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu 941  
 Db 997 GGCTCACTGTGTGACACCCACTGGGACCCAGAGAATGCCGTGTCTATGACAGACAGCTC 1056  
 Qy 942 SerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArg 961  
 Db 1057 AGCTGTGGGACTGCTCTCTCAACACAGCAGGAGAAATATATTGGAGAAAGAGTGTTCGT 1116  
 Qy 962 ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet 981  
 Db 1117 GTGGGGGACACAGGTTTCATTTGCTAGGGAATGAGTCACTTCTGGATAACTGTCAAAATG 1176  
 Qy 982 ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly 1001  
 Db 1177 ACAGTTCTTGGAGCACCTCCCTGTATCCATGGAATACTGCTCTGTGATGTCACAGGA 1236  
 Qy 1002 SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla 1021  
 Db 1237 AGCCTGACCCACCCACTGTTTCCATGCTCGCAAAATGATCTGACCCCATATTGTTCTGCA 1296  
 Qy 1022 ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly 1041  
 Db 1297 GTTCCAGAGGGCAGTGTCTTGTATCTGCTTAGAGGACAAACGGCTCCGCTAGTGTGGTGG 1356

QY 1042 AspSerArgCysAlaGlyArgValGluLeuThrHisAspGlyPheThrPglYThrIleCys 1061  
DB 1357 GACACCGCTGTGCGGAGAGTATGATCATCAGCGGCTTCTGGGCGACCATCTGT 1416  
QY 1062 AspAspGlyThrAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal 1081  
DB 1417 GATGACGGCTGGGACCTGAGCGATGCCACGTTGGTGTCAAAAGCTGGGCTGTGGAGTG 1476  
QY 1082 AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleThrPleuAsp 1101  
DB 1477 GCCTTCAATGCCACGGTCTGCTCCTTGGGAGGGGTCTAGGCCCATCTGGCTGGAT 1536  
QY 1102 AspLeuAsnCysThrGlyThrGluSerHisLeuThrPglNcysProSerArgGlyThrPglY 1121  
DB 1537 GACCTGAATGCACAGGAATGAGTCCACATGTGCGCAGTCCCTTCCCGGCGTGGGG 1596  
QY 1122 GlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeu 1141  
DB 1597 CAGCAGGACTGCAGGCACAAGGAGGACGAGGGGTCTCTGCTCAGAATTCACAGCCTG 1656  
QY 1142 ArgLeuThrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheThrAsn 1161  
DB 1657 AGGCTCTACAGTGAACCTGAACACAGAGCTGTGCTGGGAGATTGGAAGTCTTCTATAC 1716  
QY 1162 GlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCys 1181  
DB 1717 GGGACCTGGGCGAGGCTGGCAGGAGGAGAACATCACACAGCCATACGAGGATTCGTGC 1776  
QY 1182 ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGly 1201  
DB 1777 AGGCAGCTGGGCTGTGGGAGGAATGGAGTTGTACGCCCTGCGCCCTTATCTAAGACAGC 1836  
QY 1202 SerGlyPheMetTrpValAspIleGlnCysProLysThrHisIleSerIleThrPglN 1221  
DB 1837 TCTGGTTTCATGTGGTGGATGACATTGAGTCTCTAAACCATATCTCCATATGGCAG 1896  
QY 1222 CysLeuSerAlaProTrpGluArgIleSerSerProAlaGluGluThrTrpIleThr 1241  
DB 1897 TGCTGTCTGCCCATGGGAGGAAGATCTCCAGCCAGCAGAGAGACCTGGATCACA 1956  
QY 1242 CysGluAspArgIleArgValArgGlyArgThrGluCysSerGlyArgValGluIle 1261  
DB 1957 TGTGAAGATAGAAATAGAGTGGTGGAGGAGACCCGAGTGTCTGGGAGAGTGGAGATC 2016  
QY 1262 TrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGlu 1281  
DB 2017 TGGCAGCGAGGCTCTGGGCGACAGTGTGTATGACTCTGGGACCTGGCCGAGCGGAA 2076  
QY 1282 ValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPhe 1301  
DB 2077 GTGGTGTGTACAGCAGTGGGCTGTGGCTGTCTGTGGTGTGGCTGTGGGAGCGCTTCGTT 2136  
QY 1302 GlyGlnGlyThrGlyThrIleThrPleuAspMetArgCysLysGlyAsnGluSerPhe 1321  
DB 2137 GCGCAGGGAACCTGGAAACCATCTGGTGGATGACATGCGGTGCAAGGAAATGAGTCATT 2196  
QY 1322 LeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAla 1341  
DB 2197 CTATGGGACTGTACGCCAACCCCTGGGACAGAGTGTGCTGGACACACAGGAAGATGCT 2256  
QY 1342 GlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAla 1361  
DB 2257 GCGGTGAGGTGTCTGGCAGTGTGTAATCACTCAATGCTCCTCCTCAGGTCATTTAGCA 2316  
QY 1362 LeuIleLeuSerSerIlePheGlyLeuLeuLeuLeuValLeuPheIlePheLeuThr 1381  
DB 2317 CTATTTTATCCAGATCTTTGGGCTCCTTCCCTGGTGTCTGTTTATCTATTTCTCAG 2376  
QY 1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgGly 1401  
DB 2377 TGGTCCGAGTTCAGAAACAAAACATCTGCCCTCAGAGTTTCAACAGAGGAGGGGT 2436

QY 1402 SerLeuGluGluAspLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHis 1421  
DB 2437 TCTCTCGAGGAGAAATTTATTCATGAGATGGAGACCTGCCTCAAGAGAGAGACCCACAT 2496  
QY 1422 GlyThrArgThrSerAspThrProAsnHisGlyCysGluAspAlaSerAspThrSer 1441  
DB 2497 GGGACAAAGAACTCAGATGACACCCCAACCACTGTTGTGAAGATGCTAGCGACATCG 2556  
QY 1442 LeuLeuGlyValLeuProAlaSerGluAlaThrLys 1453  
DB 2557 CTGTGGGAGTCTTCTCTGCCTCTGAAGCCACAAAA 2592  
RESULT 5  
AAI58649  
ID AAI58649 standard; cDNA; 3716 BP.  
XX  
AC AAI58649;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 852.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI; 2001-442253/47.  
DR P-PSDB; AAM39493.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
XX  
PS Claim 1; SEQ ID NO 852; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and

## CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 3716 BP; 993 A; 730 C; 1060 G; 933 T; 0 other;

## Alignment Scores:

Pred. No.: 1,2e-210 Length: 3716  
 Score: 3429.00 Matches: 598  
 Percent Similarity: 71.27% Conserved: 161  
 Best Local Similarity: 56.15% Mismatches: 294  
 Query Match: 42.14% Indels: 12  
 DB: 22 Gaps: 4

US-09-759-130B-381 (1-1453) x AA158649 (1-3716)

Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383  
 Db 222 GGAACAGACAAGAGCTGAGGCTAGTGGATGGTGAACAACAAAGTGTAGCGGAGAGTGGAA 281  
 Qy 384 ValArgIleHisGluGlnTrpTrpIleCysAspGlnAsnTrpLysAsnGluAla 403  
 Db 282 GTGAAGTCCAGGAGGAGTGGGACCGGTGTATAATAGCTGGACCATGGAACGGCTC 341  
 Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys 423  
 Db 342 TCTGTGATTTGTAACAGCTGGGATGTCACACTGCTATCAAGCCCTGGATGGCTAAT 401  
 Qy 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443  
 Db 402 TCCAGTGCAGGTTCTGACCGCATTTGGATGGATCATGTTCTTCTGCGGGAATGAGTCA 461  
 Qy 444 AlaLeuTrpAspCysThrTyTrpAspGlyLysAlaLysArgTrp---CysPheArgArgSer 462  
 Db 462 GCTCTTGGGATGCAACATGATGATGGATGGGAAAGCATAGTAAGTACTGACCAACAA 521  
 Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482  
 Db 522 GATGCTGGAGTGACCTGCTCAGATGGATGCCAATTTGGAAATGAGGCTGACCGTGGAGG 581  
 Qy 483 SerProCysTyTrpGlyArgLeuGluValLysTyTrpGlnGlyGluTrpGlyValCysHis 502  
 Db 582 AATATGTTCTGGAAGAAATAGATCAAAATCCAGGACCGGTGGGAAACAGTGTGTGAT 641  
 Qy 503 AspArgTrpSerTrpArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522  
 Db 642 GATAACTCAACATAGATCATGCTCATTTGTAGACAACTGAATGTGGAAGTCT 701  
 Qy 523 MetHisValPheGlyMetThrTyTrpPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542  
 Db 702 GTCAGTTTCTCTGTTTCATCTAATTTTGGAGAAGGCTCTGACCAATCTGCTTGTGATGAT 761  
 Qy 543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562  
 Db 762 CTTATATGCAACGGAATGATGACTGCTCTGGAACCTGCAACATCAAGGATGGGAAAG 821  
 Qy 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582  
 Db 822 CATAACTGTGATCATGCTGAGGATGCTCGAGTATTGCTCAAGGGAGCATCTGAGC 881  
 Qy 583 LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyTrpPheGlnGly 602  
 Db 882 CTGAGACTGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941  
 Qy 603 ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSer 622  
 Db 942 GAATGGGGGACAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001  
 Qy 623 GlnLeuAspCysProSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyTrp 642  
 Db 1002 CACTGGGATGTCACACTGCGCTCACAGCCATGTCGATGATGATGATGATGATGATGATGAT 1061  
 Qy 643 GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys 662

Db 1062 GGACACATCTGCCTGACAGCGTTTCTGCCAGGACATGAACCTGCTGTGCGCAATGT 1121  
 Qy 663 ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys 682  
 Db 1122 AACACCATGAATGGGAAAGCATATTATTCGAATCAATCAATGAAGATGCTGGCGTGACATGT 1181  
 Qy 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702  
 Db 1182 TCTGATGATCAGATCTGGAGCTAAGACTTAGAGGTGGAGCAGCGCTGTGCTGGGACA 1241  
 Qy 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722  
 Db 1242 GTTCAGGTGGAGATTACAGACTGTTAGGGAAGTGTGTGACAGAGCGTGGGACTGAAA 1301  
 Qy 723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742  
 Db 1302 GAAGCTGATGTTGGTTGAGGAGCTGGGATGGATGTCACATCAAAACATCTTATCAA 1361  
 Qy 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762  
 Db 1362 GTGTACTCCAAATCCAGGCAACAAACACATGGCTGTTCTTAAGTAGCTGTAAACGGAAT 1421  
 Qy 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782  
 Db 1422 GAAACTTCTCTTTGGGACTGCAAGAACTGGCAATGGGTGGACTTACCTGTGATCCTAT 1481  
 Qy 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802  
 Db 1482 GAAGAAGCAAAATTACCTGCTCAGCCCAAGGAAACCCAGACTGGTGGAGGGACAAAT 1541  
 Qy 803 ProCysSerGlyArgValGluValLysHisAlaAspTrpTrpArgSerValCysAspSer 822  
 Db 1542 CCCTGTTCTGACGCTGTTGAAGTGAAGCATGTTGACACGTGGGCTCCATCTGTGATTCG 1601  
 Qy 823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842  
 Db 1602 CACTTCTCTCTGGAAGCTGCCAGCGTTCTATGCAAGGAAATACAGTGTGGCACAGTTC 1661  
 Qy 843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862  
 Db 1662 TCTATCTCGGGGAGCTCACTTTGGAGAGGAAATGGACAGACTGGGCTGGAAGAATTC 1721  
 Qy 863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882  
 Db 1722 CAGTGTGAGGACATGACTCCATCTTCACTCTGCCAGTAGCACCCGCCCAAGAAGA 1781  
 Qy 883 ThrCysIleHisSerArgGluValGlyValCysSerArgTyTrpThrAspValArgLeu 902  
 Db 1782 ACTTGTAGCCACAGCAGGAGTGTGGAGTAGTCTGCTCAAGATACACAGAATTCGCTTG 1841  
 Qy 903 ValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGly 922  
 Db 1842 GTGAATGCAAGACCCCTGTGAGGAGAGTGGAGCTCAAAACCTTGGTGCCTGGGGA 1901  
 Qy 923 SerLeuCysAspTrpHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSer 942  
 Db 1902 TCCCTCTCTAACTCTCACTGGACATAGAAGATGCCATGTTCTTTTGGCAGCAGCTAAA 1961  
 Qy 943 CysGlyThrAlaLeuSerThrThrGlyGlyLysTyTrpIleGlyGluArgSerValArgVal 962  
 Db 1962 TGTGAGTGTGCCCTTCTACCCAGGAGGACACCGTTTGGAAAAGGAATGTCAGATC 2021  
 Qy 963 TrpGlyHisArgPheHisCysLeuGluSerLeuLeuAspAsnCysGlnMetThr 982  
 Db 2022 TGGAGGATATGTTTCACTGCTGAGTGGAGCAGACATGGGAGATGTCCTTAAT 2081  
 Qy 983 ValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySer 1002  
 Db 2082 GCTCTAGTGTCTTATGATGCTTCCAGGCAAGTGGCTCTGTATCTGTCTCAGGAAC 2141  
 Qy 1003 LeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyTrpLeuSerAlaVal 1022

Db 2142 CAGTCCCAACACTGCTCTGCTGCAATTCATGCTTTGGGCCCAACAGGCGCTACCAT 2201  
Qy 1023 ProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAsp 1042  
Db 2202 CCAGAAGAAGTGTGTGGCCCTGCATAGAGAGTGTCACTTCCGCTGTGAATAGAGA 2261  
Qy 1043 SerArgCysAlaGlyArgValGluIleThrHisAspGlyPheTrpGlyThrIleCysAsp 1062  
Db 2262 GGTGCTGTGCTGGGAGACTAGATCATCATGAGGGCTCTGGGCGACCATCTGTGAT 2321  
Qy 1063 AspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAla 1082  
Db 2322 GACAGCTGGGACCTGATGATGCCACGTGGTGTTCAGACAGCTGGGCTGTGGAGAGCC 2381  
Qy 1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspAsp 1102  
Db 2382 ATTAATGCCACGTGTCTCTCATTTTGGGAAGGACAGGCGCCCATCTGGCTGGATGAG 2441  
Qy 1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGln 1122  
Db 2442 ATGAATGCAATGGAAGAAATCCCGCATTTGGCAGTGCCATTCACACGGCTGGGGCAG 2501  
Qy 1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142  
Db 2502 CAAATGTGAGGCACAAAGAGATGCGGAGTATCTGCTCAAAATCATGTCTCTGAGA 2561  
Qy 1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162  
Db 2562 CTACACAGTGAACCCAGCAGAGAGGCTGTGACGGGCGCTGGAAGTTTTTACATGGA 2621  
Qy 1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182  
Db 2622 GCITGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTGTGGTGTGGTGTGCAGG 2681  
Qy 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202  
Db 2682 CAGCTGGGCTGTGCACAAAGGAAATCAACCCCTGCATCTTTAGACAGGCGCATGTCC 2741  
Qy 1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222  
Db 2742 ATCCCATGTGGGTGGACAAATGTTAGTGTCCAAAGGACCTGACACGCTGTGGCAGTGC 2801  
Qy 1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242  
Db 2802 CCATCATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGACCTGGATCATGT 2861  
Qy 1243 GluAspArgIleArgValArgGlyAspThrGluCysSerGlyArgValGluIleTrp 1262  
Db 2862 GACAACAAGATAAGACTTCAGGAGAGACCCACTTCCTGTTGGACGTGTGGAGATCTGG 2921  
Qy 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluVal 1282  
Db 2922 CATGGAGTCTCTGGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGTCAGGTG 2981  
Qy 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302  
Db 2982 GTGTGTCAACAACACTGTGCTGTGCTCAGCTTTGAAAGCATTTCAAAAGCAGAGTGTGT 3041  
Qy 1303 GlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeu 1322  
Db 3042 CAGGGAGTCCGACCATATGCTCATGAACTGAAGTGAAGGAAATGATGATCTCTCTTC 3101  
Qy 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342  
Db 3102 TGGGATTTGCTGCCAGACGCTGGGGCCATAGTGTGAGTGTGGGCACAAAGGAGACGCTGCA 3161  
Qy 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356  
Db 3162 GTGAATTCACAGATATTTTCAGTGCAGAGAAACCCCAAAAGCCACACAGGTCGCTCA 3221  
Qy 1357 SerGlyHisLeuAlaLeuLeuSerSerIlePheGlyLeuLeuLeuValLeuPhe 1376  
Db 3222 TCCCGTCAGTCATCCTTTATTTGTCAGTCGGGATCTCTGGGGTGTCTGTGTGGCCATTTTC 3281

Qy 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394  
Db 3282 GTCGCATTATTCTTCTTCTGACTAAAAAGCGACAGACAGACGCGG-----CTTGCA 3332  
Qy 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414  
Db 3333 GTTTCCTCAAGAGGAGAGAACTTAGTCCACCAATTCATACCGGGAGATGAATCTTGC 3392  
Qy 1415 LeuLysArgGluAsp 1419  
Db 3393 CTGAATGCAGATGAT 3407  
RESULT 6  
ID AAH98740 standard; cDNA; 3834 BP.  
XX AAH98740;  
XX 12-OCT-2001 (first entry)  
XX Human EST-derived coding sequence SEQ ID NO: 597.  
DE Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition; ss.  
OS Homo sapiens.  
XX WO200154477-A2.  
PN 02-AUG-2001.  
XX 25-JAN-2001; 2001WO-US02687.  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX WPI; 2001-476164/51.  
DR P-PSDB; AAM24081.  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
PS Claim 1; Page 584-585; 1275pp; English.  
XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
CC of the invention.  
XX Sequence 3834 BP; 1017 A; 756 C; 1082 G; 979 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 1,25e-210 Length: 3834  
Score: 3429.00 Matches: 598  
Percent Similarity: 71.27% Conservative: 161  
Best Local Similarity: 56.15% Mismatches: 294  
Query Match: 42.14% Indels: 12  
Gaps: 4  
DB:

Qy	364	GlyAlaAspLeuLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu	383
Db	222	GGACAGACAGAGCGTAGCGCTAGCTGGATGGTGAACAAAGCTGTACGGGAGAGTGAA	281
Qy	384	ValArgLHisGluGlnTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAla	403
Db	282	GTCAAGAGTCCAGGAGGATGGGACCGGTGTATATATGCTGGACATGAAGACCGGTC	341
Qy	404	LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgAlaLys	423
Db	342	TCGTGATTGTAAACAGCTGGGATGCTCAACTGCTATCAAGACCCCTGGATGGCTTAAT	401
Qy	424	ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer	443
Db	402	TCCAGTGCAGGTTCTCGGACCATTTGGATGGATCATGTCTTTCTGCTGGGAATGAGTCA	461
Qy	444	AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer	462
Db	462	GCTCTTTGGGATTGCAAAACATGATGGATGGGGAAAGCATAGTAATCTGTACTACCAACAA	521
Qy	463	AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuValGlyAlaHis	482
Db	522	GATGCTGGAGTGACCTGCTACAGATGATCCCAATTTGGAATATGAGGCTGACCGCTGGAGGG	581
Qy	483	SerProCysTrpGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis	502
Db	582	AATATGTGTTCTGGAAGATAGAGATCAAAATCCCAAGCACGCTGGGAAACAGTGTGTAT	641
Qy	503	AspArgTrpSerThrArgAsnAlaAlaValCysLysGlnLeuGlyCysGlyLysPro	522
Db	642	GATAACTTCAACATAGATCATGCTGTGCTATTCTGTAGACAACATGTAATGTGGAAGTGCT	701
Qy	523	MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp	542
Db	702	GTCACTTCTCTGTTTCATCTAAATTTGGAGAGGCTCTGGACCAATCTGGTTTGATGAT	761
Qy	543	ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys	562
Db	762	CTTATATGCAACGGAATGAGTCAGCTCTCTGGAATGCAACATCAAGGATGGGAAAG	821
Qy	563	HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly	582
Db	822	CATAACTGTGATCATGCTGAGATGCTGGAGTGATTTGCTCAAGGGAGGACAGATCTGAGC	881
Qy	583	LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly	602
Db	882	CTGAGACTGTGATGCGAGTCACTGAATGTTTCAGGAAGATTAGAAGTGAGATTCCAAGGA	941
Qy	603	ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSer	622
Db	942	GAATGGGGACAATATGTATCAGCGCTGGGACATGTACGATGCTGTCTGGCATGCAAG	1001
Qy	623	GlnLeuAspCysProSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr	642
Db	1002	CAACTGGGATGCTCCACTCGGCTCAGACCATTTGGTCGAGTTAAGCCAGTAAAGGATTT	1061
Qy	643	GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys	662
Db	1062	GGACACATCTGGCTGACACGCTTTCTTGCCAGGACATGAACTGCTGTCTGGCAATGT	1121
Qy	663	ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys	682
Db	1122	AACACCATGAATGGGAAAGCATTTATTCGAATCACAATGAAGATGCTGGCGTGACATGT	1181
Qy	683	SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys	702
Db	1182	TCGTATGCATCAGATCTGGAGCTAAGACTTAGAGGTGGAGCGCCGCTGTCTGGGNACA	1241
Qy	703	ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn	722
Db	1242	GTTGAGCTGCAGATTCAGAGACTGTTAGGGAAGGTGTGTGACAGAGGCTGGGACTGAAA	1301

```
QY 1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleThrLeuAspasp 1102
Db 2382 ATTAATGCCACTGGTCTCTCTCATTTTGGGAAGAACAGGCCCATCTGGCTGGATGAG 2441
QY 1103 LeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGln 1122
Db 2442 ATGAATGCAATGGAAAGAAATCCCGCATTTGGCAGTGCATTCACAGCGCTGGGGCAG 2501
QY 1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142
Db 2502 CAAATTTGAGGACACAAAGAGATGGGGAGTTATCTGCTCAGAAATTCATCTCTGAGA 2561
QY 1143 LeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValaPheTyrAsnGly 1162
Db 2562 CTGACACGTGAAGCCACGACAGAGAGCGCTGCGAGGCGCTCTGGAAGTTTTTACATGGA 2621
QY 1163 ThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCysArg 1182
Db 2622 GCTTGGGCGACTGTGGCAAGAGTAGCATGCTGAAACCACATGTGGGTGGTGTGGCAGG 2681
QY 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202
Db 2682 CAGCTGGCTGTGCAGACAAAGGAAATCAACCTGTCATCTTAGACAAAGCCCATGTCC 2741
QY 1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCys 1222
Db 2742 ATTCATCTCCATGGTGGTGGACATGTTTCAGTGTCCAAAGGACCTGCACAGCTGTGGCAGTCC 2801
QY 1223 LeuSerAlaProTrpGluArgGlySerProAlaGluThrTrpIleThrCys 1242
Db 2802 CCATCATCTCCATGGGAGAGAGAGCTGGCCAGCCCTCGAGAGAGACCTGGATCACATGT 2861
QY 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262
Db 2862 GACAACAAGATAAGACTTCAGGAAGAGACCCACTTCCTGTTCTGGAGCTGTGGAGATCTCG 2921
QY 1263 HisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluVal 1282
Db 2922 CATGGAGTTCTCGGGGACAGTGTGTGATGACTCTTGGGACTTGGACGATGCTCAGGTG 2981
QY 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGly 1302
Db 2982 GTGTGTCAACAACCTGGCTGTGTCCAGCTTTGAAGCATTTCAAGAAGCAGAGTTTGT 3041
QY 1303 GlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeu 1322
Db 3042 CAGGGGACTGGACCGATATGGCTCAATGAAGTGAAGTCAAAAGGAATGAGTCTTCTCTTG 3101
QY 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342
Db 3102 TGGGATGTCTCGCCAGACGCTGGGGCCATAGTGTGTGGCAGCAAGCAAGCAGCTGCA 3161
QY 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356
Db 3162 GTGAATTCACAGATATTTCACTGTCAGTGCAGAAACCCCAAAAGCCACACAGCTCCTCA 3221
QY 1357 SerGlyHisLeuAlaLeuIleLeuSerIlePheGlyLeuLeuLeuValLeuPhe 1376
Db 3222 TCCCGTCAGTCATCTTTATTCAGTCGGGATCCCTTGGGGTGTCTGTGTGGCCATTTTC 3281
QY 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394
Db 3282 GTCCGATTTATCTCTGTACTAAAGCAAGCAAGCAGACAGACGG-----CTTGCA 3332
QY 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414
Db 3333 GTTTCCTCAAGAGAGAGAACTTAGTCCACCAAAATCAATACCGGGAGATGAATCTTCTTC 3392
QY 1415 LeuLysArgGluAsp 1419
Db 3393 CTGAATGCAGATGAT 3407
RESULT 7
```

```
AAI60435/c
ID AAI60435 standard; cDNA; 3834 BP.
XX
AC AAI60435;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4424.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 13-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
(PHYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Dmanac RT;
WPI; 2001-442253/47.
DR P-PSDB; AAM41279.
DR
PT Novel nucleic acids and polypeptides, useful for treating disorders
such as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 4424; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
the encoded polypeptides (AAM38642-AAM42213) with nootropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
specification.
XX
SQ Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1,25e-210 Length: 3834
Score: 3429.00 Matches: 598
Percent Similarity: 71.27% Conservative: 161
Best Local Similarity: 56.15% Mismatches: 294
Query Match: 42.14% Indels: 12
DB: 22 Gaps: 4
```



QY 1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleThrLeuAspAsp 1102  
DB 1453 ATTAATGCCACTGGTCTCTCATTTGGGGAAGGACAGGCCCATCTGGCTGGATGAG 1394  
QY 1103 LeuAsnCysThrGlyThrGluSerHisLeuThrGlnCysProSerArgGlyTrpGlyGln 1122  
DB 1393 ATGAATGCAATGAAAGAAATCCCGCATTTGGCAGTCATTCACAGCGCTGGGGGAG 1334  
QY 1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142  
DB 1333 CAAATTTGAGGACAGGAGATGGGGAGTATCTGCTAGAATTCATGTCCTGAGA 1274  
QY 1143 LeuTyrSerGlyThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGly 1162  
DB 1273 CTGACCAAGTGAACCCAGCAGAGAGCGCTGTGAGGCGCTGGAAGTTTTTACAATGCA 1214  
QY 1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182  
DB 1213 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAACCACTGTGGGTGTGTGTCAGG 1154  
QY 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202  
DB 1153 CAGCTGGGTGTCACACAAAGGAAATCAACCTGTCATCTTAGACAGGCCATGTCC 1094  
QY 1203 GlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleThrGlnCys 1222  
DB 1093 ATTCCCATGTGGTGACAAATGTTAGTCTCCAAAGGACCTGACACCGCTGTGGCAGTGC 1034  
QY 1223 LeuSerAlaProTrpGluArgArgIleSerSerProAlaGluThrTrpIleThrCys 1242  
DB 1033 CCATCATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGACCTGGATCATGTT 974  
QY 1243 GluAspArgIleArgValArgGlyValAspThrGluCysSerGlyArgValGluIleTrp 1262  
DB 973 GACAAACAGATAGACTTCAGGAGGACCCACTCTCTGTCGAGCTGGGAGATCTGG 914  
QY 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluVal 1282  
DB 913 CATGGAGTTCCTGGGGGACAGTGTGTGATGACTCTGGGACTGGACGATGCTCAGGTG 854  
QY 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGly 1302  
DB 853 GTGTGCAACAACTTGCTGTGCTGCTGAGCTTTGAAGCAATTCAAAGACAGAGTGTGT 794  
QY 1303 GlnGlyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeu 1322  
DB 793 CAGGGGACTGGACCATATGGCTCAATGAAGTGAAGTCAAGGGAATGAGTCTCTCTTG 734  
QY 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342  
DB 733 TGGGATTGTCTGCCAGACCTGGGGCCATAGTGTGGGACCAAGGAAGACGCTGCA 674  
QY 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356  
DB 673 GTGAATGTCACAGATATTTCAGTGCAGAAACCCCAACAAACCCACAGCTCCCTCA 614  
QY 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuValLeuPhe 1376  
DB 613 TCCCGTCAGTCATCTTTATTGTCAGTCGGGATCCTTGGGGTGTCTGTGTGGCCATTTC 554  
QY 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394  
DB 553 GTCGCAFTATTCTTCTGATAAAAGCGAAGACAGACAGACGCG-----CTTGCA 503  
QY 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414  
DB 502 GTTCTCCAGAGAGAGAACTTAGTCCACCAATTCATACCGGAGATGAATCTTTC 443  
QY 1415 LeuLysArgGluAsp 1419  
DB 442 CTGAATCCAGATGAT 428

RESULT 8  
AAI60436/c  
ID AAI60436 standard; cDNA; 3834 BP.  
XX  
AC AAI60436;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 4425.  
XX  
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
peripheral nervous system; neuropathy; central nervous system; CNS;  
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukaemia; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
(HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI; 2001-442253/47.  
DR P-PSDB; AAM41280.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
such as central nervous system injuries -  
XX  
PS Claim 1; SEQ ID NO 4425; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
immunosuppressant and cytostatic activity. The polynucleotides are useful  
in gene therapy. A composition containing a polypeptide or polynucleotide  
of the invention may be used to treat diseases of the peripheral nervous  
system, such as peripheral nervous injuries, peripheral neuropathy and  
localised neuropathies and central nervous system diseases, such as  
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
assays for receptor activity, arthritis and inflammation, leukaemias and  
C.N.S disorders.  
CC Note: the sequence data for this patent did not form part of the printed  
specification.  
XX  
SQ Sequence 3834 BP; 979 A; 1082 C; 756 G; 1017 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 1.25e-210 Length: 3834  
Score: 3429.00 Matches: 598  
Percent Similarity: 71.27% Conservatives: 161  
Best Local Similarity: 56.15% Mismatches: 294  
Query Match: 42.14% Indels: 12

DB:	22	Gaps:	4
US-09-759-130B-381 (1-1453) x AA160436 (1-3834)			
Qy	364	GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu	383
Db	3613	GGAAACAGACAAGAGCTGAGGCTAGTGATGTTGGAACAACAGTAGTCGGGAGAGTAGTGAA	3554
Qy	384	ValArgIleHisGlnIlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlnAla	403
Db	3553	GTGAAGATCCAGGAGGAGTGGGAACGGTGTGTAAATAAGTGTGAGCATGTGAAGCGGTC	3494
Qy	404	LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys	423
Db	3493	TCGTGATTTGAACCATCGGGATGTCACACTATCAAAAGCCCTGGATGGCGTAAAT	3434
Qy	424	ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer	443
Db	3433	TCAGATGCAGGTTCTCGACGCATTTGGATGGATCATGTTCTTGTCTGGGAATGAGTCA	3374
Qy	444	AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer	462
Db	3373	GCCTCTTGGGATTGCAACATCATGATGGGGGAAACATAGTAAGTGTACTACCAACAA	3314
Qy	463	AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis	482
Db	3313	GATGCTGGAGTGACCTGCTCAGATGGATCCCAATTTGGAAATCAGCGTCAGCGGTGAGG	3254
Qy	483	SerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis	502
Db	3253	ANTATGTGTTCTGGAAGAATAGAGATCAAAATTCCAAGGACCGTGGGAAACAGTGTGTGAT	3194
Qy	503	AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro	522
Db	3193	GATAACTTCAACATAGATCATGCATCTGTCTATTCTGACAACTGAATGTGGAAGTCT	3134
Qy	523	MetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp	542
Db	3133	GTCAGTTTTCTCTGGTTTCATCTAATTTGGAGAAAGCTCTGGACCAATCTGCTTTGATGAT	3074
Qy	543	ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys	562
Db	3073	CTTATATGCAACGGAAATGAGTCAGCTCTCTGGAACTGCAAAACATCAAGGATGGGAAAG	3014
Qy	563	HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly	582
Db	3013	CATAACTGTGATCATCTCTGAGGATGCTGGAGTGATTGCTCAAGGGAGCAGATCTGAGC	2954
Qy	583	LeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly	602
Db	2953	CTGAGACTGTGATAGATGGAGTCACTGAATGTTTCAGAAAGATTAGAAAGTGAAGATCCAAAGGA	2894
Qy	603	ArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSer	622
Db	2893	GAATGGGGACAATATGTATGACGCTGGACAGTTACGATGCTGCTGGCATGCAAG	2834
Qy	623	GlnLeuAspCysProSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr	642
Db	2833	CAACTGGGATGTCCTCACTCCGCTCAGACCATTTGGTCGAGTTACGCCAGTAAAGGATTT	2774
Qy	643	GlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCys	662
Db	2773	GGACACATCTGGCTTGACACCGTTTCTTGCCAGGCACATGAACCTGTCTCTGCGCAATGT	2714
Qy	663	ArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCys	682
Db	2713	AAACACCATGAATGGGGAAGCATTTATTGCAATCAATGAAGATGCTGGCGTGACATGT	2654
Qy	683	SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys	702
Db	2653	TCGTGATGGATCAGATCTGGAGCTTAGACTTTAGAGGTGGAGCGCCGCTGTCTGGGACA	2594
Qy	703	ValcluuValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn	722

Db 1513 GACAGCTGGGACCTGAGTATGCCACCTGGTGGTTGTCAGACACACTGGGCTGTGAGAGGCC 1454  
QY 1083 PheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleThrLeuAspArg 1102  
Db 1453 ATTAATGCCACTGGTCTCTCTCATTTTGGGAAGGAACAGGCCCATCTGGCTGGATGAG 1394  
QY 1103 LeuAsnCysThrGlyThrGluSerHisLeuThrPheGlnCysProSerArgGlyTrpGlyGln 1122  
Db 1393 ATGAATGCAATGGAAGAAATCCCGCATTTGGCATGTCATTCACACCGCTGGGGGCGAG 1334  
QY 1123 HisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArg 1142  
Db 1333 CAATATGTCAGGACAGGAGGATGCGGAGGTATCTGCTCAGAAATCATGTCTCTGAGA 1274  
QY 1143 LeuThrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluValPheThrAsnGly 1162  
Db 1273 CTGACCAAGTGAAGCCAGCAGAGAGAGCGCTGTGCAGGCGCTCTGGAAGTTTTTACAATGGA 1214  
QY 1163 ThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArg 1182  
Db 1213 GCTTGGGGCACTGTTGGCAAGAGTAGCATGTCTGAAACCACTCTGGGTGTGTGTGCAGG 1154  
QY 1183 GlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySer 1202  
Db 1153 CAGCTGGGCTGTCAGCAAGAAAGGAAATCAACCTGTCATCTTTAGCAAGGCCATGTCC 1094  
QY 1203 GlyPheMetTrpValAspAlaGlnCysProLysThrHisLeuSerIleThrPheGlnCys 1222  
Db 1093 ATTCCTCATGTTGGTGACATGTTCACTGTCCTCAAAAGGACCTGACACGCTGTGGCAGTGC 1034  
QY 1223 LeuSerAlaProTrpGluArgIleSerSerProAlaGluGluThrTrpIleThrCys 1242  
Db 1033 CCATCATCTCCATGGGAGAGAGACTGGCCAGCCCTCGGAGGAGACCTGGATCATGT 974  
QY 1243 GluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrp 1262  
Db 973 GACACCAAGATAGACTCAGGAGGACCCACTTCTGTCGAGCTGTGGAGATCTGG 914  
QY 1263 HisAlaGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaVal 1282  
Db 913 CATGGAGGTTCTTGGGGACAGTGTGTGATGACTCTGGGACTTGGACGATGCTCAGGTG 854  
QY 1283 ValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGly 1302  
Db 853 GTGTGTCAACACTGGCTGTGGTCCAGCTTTGAAGCATTCAAAGACGATGAGTTGT 794  
QY 1303 GlnClyThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeu 1322  
Db 793 CAGGGACTGGACCATATGCTCAATGAGTGAAGTGAAGGGAATGAGTCTCTCTTG 734  
QY 1323 TrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGly 1342  
Db 733 TGGGATTGTCTCCAGACGCTGGGCCCATAGTGTGGGCACAAAGAACGCGTGCA 674  
QY 1343 ValArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSer 1356  
Db 673 GTAATGTGCACAGATATTTCACTGTCAGAAACCCCAACAAAGCCACACAGCTCCCTCA 614  
QY 1357 SerGlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuValLeuPhe 1376  
Db 613 TCCCGTCAGTCATCCTTTATTCAGTCGGGATCCTTGGGGTGTCTGTGGCCATTTTC 554  
QY 1377 Ile-----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArg 1394  
Db 553 GTGGCATTTATCTCTGTGACTAAAGGCAAGACAGACAGCGG-----CTTGCA 503  
QY 1395 ValSerThrArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCys 1414  
Db 502 GTTCTCTCAAGGAGGAGAGAACTTAGTCCACCAAAATCAATACCGGGAGATGAATTTCTGC 443  
QY 1415 LeuLysArgGluAsp 1419  
Db 442 CTGAATGCAGATGAT 428

## RESULT 9

AAI58650  
ID AAI58650 standard; cDNA; 3811 BP.  
XX  
AC AAI58650;  
XX 22-OCT-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 853.  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX  
OS Homo sapiens.  
PN WO200153312-A1.  
XX 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US34263.  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
DR P-PSDB; AAM39494.  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
PS Claim 1: SEQ ID NO 853; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX Sequence 3811 BP; 1018 A; 751 C; 1081 G; 961 T; 0 other;

## Alignment Scores:

Pred. NO.: 1.51e-208 Length: 3811  
Score: 3396.50 Matches: 597  
Percent Similarity: 69.03% Conservative: 161  
Best Local Similarity: 54.37% Mismatches: 295

```

Query Match: 41.74% Indels: 45
DB: 22 Gaps: 5
US-09-759-130B-381 (1-1453) x AA158650 (1-3811)

Qy 364 GlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArgValGlu 383
Db 222 GGAACAGACAGAGAGCTAGCTAGTGGATGGTGAAGAAAGTGTAGCGGAGAGTGGAA 281
Qy 384 ValArgIleHisGluGlnTrpThrIleCysAspGlnAsnTrpLysAsnGluAla 403
Db 282 GTGAAGTCCAGGAGGAGTGGGAAACGCTGTGTAATATGGCTGGAGCATGGAACGGCTC 341
Qy 404 LeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgAlaLys 423
Db 342 TCTGTGATTGTACACGCTGGGATGTCACACTGCTATCAAGCCCTGGATGGCTAAT 401
Qy 424 ProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSer 443
Db 402 TCCAGTCGAGGTTCTGGACGCATTTGGATGGATCATGTTCTTCTGCTGGGAATCAGTCA 461
Qy 444 AlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThr---CysPheArgArgSer 462
Db 462 GCTCTTTGGGATGCAACATGATGGATGGGAAAGCATAGTAACCTGTACTCACCACAA 521
Qy 463 AspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHis 482
Db 522 GATGCTGGAGTGACCTGCTCAGATGGATCCCAATTTGGAATGAGGCTGACCCGCTGGAGG 581
Qy 483 SerProCysThrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHis 502
Db 582 AATATGTGTTCTGGAAGAATAGATCAAAATTCGAAGACGCTGGGAAACATGTGTGAT 641
Qy 503 AspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPro 522
Db 642 GATAACTTCACATAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
Qy 523 MethisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAspAsp 542
Db 702 GTCAGTTTCTGCTGCTCAATTTTGGAGAAGGCTCTGGACCAATCTGCTGCTGCTGCTGCT 761
Qy 543 ValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLys 562
Db 762 CTTATATGCAACGGAATAGCTAGCTCTCTGGAAGTCAACATCAAGAGTGGGGAAG 821
Qy 563 HisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGly 582
Db 822 CATACTGTGATCATGCTGAGATGCTGAGTGTGCTCAAGGGAGCAGATCTGAGC 881
Qy 583 LeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGly 602
Db 882 CTGAGACTGGTAGATGGATCACTGAATGTTCCAGGAAGATAGAGTGAGATTCGAAGGA 941
Qy 603 ArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCysSer 622
Db 942 GAATGGGGGACAATATGATGACGCTGGGACAGTACGATGCTGCTGCTGCTGCTGCTGCTGCT 1001
Qy 623 GlnLeuAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyr 642
Db 1002 CAACTGGGATGTCACACTGCGCTCACACCCATGCTGAGTAAACCCAGTAAGGGATTT 1061
Qy 643 GlyLysIleTrpLeuAspValPheSerCysAspGlyAspGluSerAspLeuTrpSerCys 662
Db 1062 GGACACATCTGGCTTGACAGCGTTTCTTCCAGGAGCATGAACCTGCTGCTGCTGCTGCTGCT 1121
Qy 663 ArgAsnSerGlyTrpGlyLysAsnAspCysSerHisSerGluAspValGlyValIleCys 682
Db 1122 AAACACCATGATGGGAAACATATTGCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1181
Qy 683 SerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGlyLys 702
Db 1182 TCTGATGATCAGATCTGAGACTTAAGCTTAGAGGTGGAGCGACCCGCTGTGCTGGGACA 1241

```

---

```

Qy 703 ValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMetAsn 722
Db 1242 GTTCAGGTGGAGATTACAGACTCTTAGGGAAGTGTGTGACAGAGGCTGGGACTGAAA 1301
Qy 723 IleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGlu 742
Db 1302 GAAGCTGATGTGGTTTCAGGCAGCTGGGATGTGATCTGCACATAAACAATCTTATCAA 1361
Qy 743 ProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGly 762
Db 1362 GTGTACTCCAAAATCCAGGCAACACATCGCTGTCTTCTAAGTAGCTGTAAACGGAAT 1421
Qy 763 GluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsn 782
Db 1422 GAAACTTCTCTTTGGGACTGCAAGAACTGCAATGGGGTGGACTTACCTGTGATCATTAT 1481
Qy 783 MetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAlaAspMet 802
Db 1482 GAAGAAGCCAAAATTTACCTGCTCAGCCACAGGAACCCAGACTGGTTGGAGGGGACATT 1541
Qy 803 ProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSer 822
Db 1542 CCTGTCTTGGACGTGTGAAGTGAAGCATGCTGCACAGCTGGGCTCCATCTGTGATTCTG 1601
Qy 823 AspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIle 842
Db 1602 GACTTCTCTGGAAGCTGCCAGGCTTCTATGACGGAATTTACAGTGTGGCACACTTGTCT 1661
Qy 843 SerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPhe 862
Db 1662 TCTATCTCTGGGGAGCTCACTTTGGAGAGGAAATGACAGATCTGGGCTGAAGAAATTC 1721
Qy 863 GlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAsp 882
Db 1722 CAGTGTGAGGACATGATGCTCCATCTTCTCCTGCCAGTAGCACCCGCCCAAGGA 1781
Qy 883 ThrCysIleHisSerArgGluValGlyValValCysSerArg----- 896
Db 1782 ACTTGTAGCCACAGCAGGGATGTTGGAGTAGTCTCTCAAGTAAGACCCAGAAAAACATCT 1841
Qy 896 ----- 896
Db 1842 TTAATTTGTTCTTATCTGTGAAAGGACAGGTTAGGAGTCAATAGCTGTCTTTTCTTA 1901
Qy 897 -----TyrThrAspValArgLeuValAsnGlyLysSerGlnCys 909
Db 1902 AAGCCCTCTCTCTCCAGGATACACAGAAATTCCTGGTGAATGGCAGACCCCGTGT 1961
Qy 910 AspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrp 929
Db 1962 CAGGCGAGGTGGAGCTCAAAACGCTTGGTGCCTGGGATCCCTCTGTAACCTCTCACATGG 2021
Qy 930 AspProGluAspAlaArgValIleCysArgGlnLeuSerCysGlyThrAlaLeuSerThr 949
Db 2022 GACATAGAAGATGCCATGTTCTTTGCCAGCAGCTTAATATGGAGTTGCCCTTCTTACC 2081
Qy 950 ThrGlyGlyLysTyrIleGlyLysArgSerValArgValTrpGlyHisArgPheHisCys 969
Db 2082 CCAGGAGGACGACGCTTTTGGAAAGAAATGCTCAGATCTGGAGGCATATGTTTCCATGC 2141
Qy 970 LeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProCys 989
Db 2142 ACTGGAGTACGACGACATGGAGATTGCTGCTGTAACATGCTCTAGGTGCTTCATTATGT 2201
Qy 990 IleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGlnProLeuPhePro 1009
Db 2202 CCTTCAGAGCAAGTGGCTCTGTAATCTGTCAGGAACACAGTCCCAACACATGTCCTCG 2261
Qy 1010 CysLeuAlaAsnValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIle 1029
Db 2262 TGCATATCATCTCTTTGGGCCCCAACAGAGGCTACCATTCAGAGAAAGTGTGTGCC 2321
Qy 1030 CysLeuGluAspLysArgLeuValValAspGlyAspSerArgCysAlaGlyArgVal 1049

```

Db 2322 TGCATAGAGAGTGGTCAACTTCGCTGTAATGAGGAGGTGCTGCTGGGAGATA 2381  
QY 1050 GluIleThrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp 1069  
Db 2382 GAGATCTATCATGAGGAGTCCCTGGGGCCACCATCTGTGTGATGAGCTGGGACCTGAGTGAT 2441  
QY 1070 AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla 1089  
Db 2442 GCCACAGTGGTTGGACAGAGCTGGGCTGTGGAGAGGCCATTAATGCCACATGGTCTCTGCT 2501  
QY 1090 HisPheGlyGluCysGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu 1109  
Db 2502 CATTTTGGGAAAGAACAGGGCCCATCTGCTGGTGGATGAGTGAATGGAAAGAA 2561  
QY 1110 SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGlu 1129  
Db 2562 TCCCGCATTTGGCAGTGGCAATTCACACGGCTGGGGGCAACAAATTTGGAGCCACAGGAG 2621  
QY 1130 AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTrpSerGluThrGluThr 1149  
Db 2622 GATCGGGAGTTATCTGCTCAGCAATTCATGCTGTGAGACTGACCATGAGCCAGCAGAGA 2681  
QY 1150 GluSerCysAlaGlyArgLeuGluValPheThrAsnGlyThrTrpGlySerValGlyArg 1169  
Db 2682 GAGGCTGTGACAGGCGTCTGGAAGTTTTTACAAATGGAGCTTGGGGCAGCTTTGGGCAAG 2741  
QY 1170 ArgAsnIleThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn 1189  
Db 2742 AGTAGCATGCTGAAACCACTGTGGTGTGGTGTGAGGCGACCTGGGCTGTCACACAAA 2801  
QY 1190 GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp 1209  
Db 2802 GGGAAATCAACCTGCATCTTTAGACAGGCGCATGTCCATTCCTCCATGGGTGGACAAT 2861  
QY 1210 IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg 1229  
Db 2862 GTTCAGTGTCCAAAGAGGAGCTGACAGCTGTGGCAGTGGCAGTCCATCCATGGGAGAG 2921  
QY 1230 ArgIleSerProAlaGluThrTrpIleThrCysGluAspArgIleArgValArg 1249  
Db 2922 AGATGGCCAGCCCTCGGAGAGACCTGATCATCATGTGACAAACAGATAGACTTCAG 2981  
QY 1250 GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr 1269  
Db 2982 GAAGGCCACCATCTCTGTCTGGAGCTGTGGAGATCTGCATGGAGGTTCCTGGGGGACA 3041  
QY 1270 ValCysAspAspSerTrpAspLeuAlaGluAlaValValCysGlnGlnLeuGlyCys 1289  
Db 3042 GTGTGTGATGACTCTTGGGACTTGGAGTGTGAGTGTGAGTGTGATCAACAACCTTGGCTGT 3101  
QY 1290 GlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrp 1309  
Db 3102 GTGCCAGCTTGAAGACCATCAAGAGCAGAGATTTGGTCAGGGGACTGGCCGATATGG 3161  
QY 1310 LeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysPro 1329  
Db 3162 CTCATGAAGTGAAGTGAAGAGGAGATGAGTCTCTTGTGGATGTCTCCCTCCAGACGC 3221  
QY 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSer 1349  
Db 3222 TGGGGCCATAGTGTGGGACAGAGGAGCGCTGCGAGTGAATTTGCACATATTTCA 3281  
QY 1350 LeuLys-----SerLeuAsnAlaSerSerGlyHisLeuAlaLeuIle 1363  
Db 3282 GTGCAGAAAAACCCCAAAAGCCACAGAGTGGCTCATCCGTCAGTCATCTTAT 3341  
QY 1364 LeuSerSerIlePheGlyLeuLeuLeuValLeuPheIle-----LeuPheLeuThr 1381  
Db 3342 GCAGTCGGGATCTTGGGGTGTCTGTGGCCATTTTCTGCGCATTTATCTTCTTGACT 3401  
QY 1382 TrpCysArgValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgGly 1401

Db 3402 AAAAAGCGAAGACAGACAGACAGCGG-----CTTGCAAGTTTCTCAAGAGGAGAGAAC 3452  
QY 1402 SerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAsp 1419  
Db 3453 TTATCCACCAAAATTCATACCGGAGATGAATCTTCCTGCTGAATGCAGATGAT 3506  
RESULT 10  
AAS94922  
ID AAS94922 standard; DNA; 3707 BP.  
XX AAS94922;  
XX 14-FEB-2002 (first entry)  
XX Human DNA sequence #177 expressed during foam cell differentiation.  
XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;  
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.  
XX Homo sapiens.  
PN WO200177389-A2.  
XX 18-OCT-2001.  
XX 04-APR-2001; 2001WO-US11128.  
XX 05-APR-2000; 2000US-195106P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;  
PI Tai J;  
XX WPI; 2002-010925/01.  
XX Composition useful for diagnosis of conditions, disorders or diseases  
PT associated with atherosclerosis, comprises several polynucleotides that  
PT are differentially expressed in foam cell development -  
Claim 1; Page 232-233; 315pp; English.  
XX The present invention relates to the isolation of human polynucleotide  
CC sequences that are differentially expressed during foam cell  
CC differentiation. The polynucleotide sequences of the invention or a  
CC composition comprising these polynucleotides are useful as a high  
CC throughput method for detecting altered expression of one or more  
CC polynucleotides in a sample. The polynucleotides can be used in the  
CC diagnosis of disorders associated with foam cell development such as  
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as  
CC coronary artery disease. The polynucleotide sequences can also be used  
CC as PCR primers and probes. The polynucleotides of the invention are also  
CC useful in gene therapy. AAS94746-AAS95021 represent the human  
CC polynucleotide sequences of the invention which are differentially  
CC expressed during foam cell differentiation.  
XX Sequence 3707 BP; 984 A; 731 C; 1062 G; 930 T; 0 other;  
SQ

## Alignment Scores:

Pred. No.:	2,61e-207	Length:	3707
Score:	3377.00	Matches:	598
Percent Similarity:	71.00%	Conservative:	161
Best Local Similarity:	55.94%	Mismatches:	294
Query Match:	41.50%	Indels:	16
DB:	24	Gaps:	4

US-09-759-130B-381 (1-1453) x AAS94922 (1-3707)

QY 364 GlyAlaAspLeuGluArgLeuAlaAspGlySerAsnAsnCysSer-GlyArgValG 383  
Db 222 GGAACAGACAGAGAGCTGAGGTAGTGGATGGTGAACCAAGTGTAGCGGGAGAGTGA 281  
QY 383 uValArgIleHisGluGluTrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAl 403

Db	282	AGTGAAGTCCAGGAGAGTGGGAACGGTGTGTAAATAATGGCTGGAGCATGGGAAGCGGT	341
Qy	403	aLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArgAlaLys	423
Db	342	CTCTGTGATTGTACCAACGAGTGGGATGTCCTCAACTGCTATCAAGCCCTGGATGGGCTAA	401
Qy	423	sProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSe	443
Db	402	TTCCAGCTGCAGGTCTTGGAGCAGATTTGGATGGATCATGTTTCTTCTGCTGGGAATGAGT	461
Qy	443	rAlaLeuTrpAspCysThrTyArgGlyLysAlaLysArgThr--CysPheArgArgSe	462
Db	462	AGTCTTTGGGATTTGCAACATGATGGATGGGAAGAGCATAGTAACGTGTACTCACCANCA	521
Qy	462	rAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHi	482
Db	522	AGATGCTGGAGTACCTGCTCAGATGGATCCAATTTGGAATGAGGCTGACGCGTGGAGG	581
Qy	482	sSerProCysTyArgLeuGluValLysTyArgGlyClnGlyTrpGlyThrValCysHi	502
Db	582	GAATATGTGTCTTGGGAAGTAAGACATAATTCACGAGCGGTGGGAACAGCTGTGTGA	641
Qy	502	sAspAlaGlyTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLysPr	522
Db	542	TGATACTTCAACATAGATCATGCTGTGCTATTGTAGACAACTTGAATGTGGAAGTGC	701
Qy	522	oMetHisValPheGlyMetThrTyArgLysGlyAlaSerGlyProIleTrpLeuAspAs	542
Db	702	TGTCAGTTCTCTGGTTCATCTAATTTGGAGAAGCTCTGGACCAATCTGGTTTGATGA	761
Qy	542	pValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGlyLy	562
Db	762	TCTTATATGACAGGAATAGTCAAGCTCTCTGGAACATGCAACATCAAGATGGGAAA	821
Qy	562	sHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrpGI	582
Db	822	GCATACTGTGATCATGCTGGAGATGCTGGAGTGATTGCTCAAGGAGCAGCATCTGAG	881
Qy	582	yLeuArgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyPheGlnGI	602
Db	882	CCTGAGACTGGTAGAGGAGTCACTCAATGTTCAGAGAAGATTAGACTGAGATCCAAAG	941
Qy	602	yArgTrpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaLavalValCysSe	622
Db	942	AGAATGGGGACAATATGTGATCAGCGGTGGCAGACTTACGATGCTGCTGGCATGCA	1001
Qy	622	rGlnLeuAspCysProSerSerIleGlyMetClyLeuGlyAsnAlaSerThrGlyTy	642
Db	1002	GCACCTGGGATGTCCAACTGCGCGTCACAGCCATTTGGTTCAGTTCACGCGAGTAAGG	1061
Qy	642	rGlyLysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCy	662
Db	1062	TGGACACATCTGGCTTGACAGCGTTTCTTCCAGGGACATCAACTGCTGTCTGGCAATG	1121
Qy	662	sArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCy	682
Db	1122	TAAACACCATGAATGGGGAAGCATTTATTGCAATCACAATGAAGATGCTGGCGTGCATG	1181
Qy	682	sSerAspAlaSerAspMetGluValArgLeuValGlyGlySerSerArgCysAlaGlyLy	702
Db	1182	TTCTGTATGATCATGATCTGGAGCTTAAGACTTAAGAGGTGGAGCGGACGCGCTGTCTG	1241
Qy	702	sValGluValAsnValGlnGlyAlaValGlyLeuCysAlaAsnGlyTrpGlyMetAs	722
Db	1242	AGTTAGGTGGAGATTCAGAGACTGTTAGGGAAGGTGTGTACAGAGCGCTGGGACATGA	1301
Qy	722	nIleAlaGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGI	742
Db	1302	AGAAGCTGATGTGGTGTGCAGGCACCTGGGATGTGGATCTCCACTCAAAACATCTTATCA	1361
Qy	742	uProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGI	762

D	b	1362	AGTGACTCCAAATCCAGCAACAACACATGGCTGTTTCTTAAGTACTGCTGTACGGAAA	1421
Q	y	762	yGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAs	782
D	b	1422	IGAAACTTCTTTGGGACTGCAAGAACTGGCAATGGGTGGACTTACCTGTGATCACTA	1481
Q	y	782	nMetGluAlaSerLeuIleCysSer-AlaHisArg-GlnProArgLeuValGlyAlaAsp	801
D	b	1482	TGAAGAAGCCAAAATTACCTGCTCAGGCCCCACAGGGAAACCCAGACTGGTTGGAGGGAC	1541
Q	y	802	MetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAsp	821
D	b	1542	ATTCCCTGTTCTGGAGCTGTTGAAGTGAACATGGTGACAGCTGGGGCTCCATCTGTGAT	1601
Q	y	822	SerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAla	841
D	b	1602	TCAGACTTCTCTCGGAAGCTGCACGCTTCTATGCAGGGAATTACAGTGTGGCACAGTT	1661
Q	y	842	IleSerLeuSerValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLys	861
D	b	1662	GTCTCTATCTCGGGGAGGCTCACTTTGGAGAGGGAATGGACAGATCTGGGCTGAAGAA	1721
Q	y	862	PheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGlu	881
D	b	1722	TTCCAGTGTGAGGACATGATGCCCATCTTTCACCTGCCAGTAGACACCCCGCCAGAA	1781
Q	y	882	AspThrCysIleHisSerArgGluValGlyValValCysSerArgTyrThrAspValArg	901
D	b	1782	GGAACTTGTAGCCACACAGCAGGAGTGTGGAGTAGTCTGCTCAAGATACACAGAAATTCG	1841
Q	y	902	LeuValAsnGlyLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrp	921
D	b	1842	TTGTGTAAATGGCAAGACCCCGTGTGAGGCAGAGTGGAGCTCAAAACGCTTGGTCCCTGG	1901
Q	y	922	GlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeu	941
D	b	1902	GGATCCCTCTGTAACTCTACTGGGACATAGAAGATGCCCATGTTCTTTGCCAGCAGCTT	1961
Q	y	942	SerCysGlyThrAlaLeuSerThrGlyGlyLysTyrIleGlyGluArgSerValArg	961
D	b	1962	AAATGTGGAGTTGCCCTTTCTACCCAGGAGGAGCAGCGTTTGGAAAAAGAAATGGTCAG	2021
Q	y	962	ValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMet	981
D	b	2022	ATCTGGAGGATATGTTTCTACTGCATCGGAGCTGGACAGCACATGGGAGATGTCTCTGA	2081
Q	y	982	ThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCysThrGly	1001
D	b	2082	ACTGCTCTAGTGCTTCATATATGCTTCAGACCAAGTGGCCCTCTGTAACTGCTCAGGA	2141
Q	y	1002	SerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAla	1021
D	b	2142	AACCAAGTCCCAACACATGTCCTCGTGCANTTCATGCTCTTTGGGCCCAACAGGCCCTACC	2201
Q	y	1022	ValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGly	1041
D	b	2202	ATTCACAGAAGAGTGTGTGGCCTGCATAGAGAGTGGTCAACTTCTGCCTGCTGTAATGGA	2261
Q	y	1042	AspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCys	1061
D	b	2262	GGAGGTCCGTGCTCTGGAGAGTAGAGATCTATCATGAGGGCTCTCTGGGGCACCATTCTGT	2321
Q	y	1062	AspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyVal	1081
D	b	2322	GATGACAGCTGGGACCTGAGTGATCCCACTGGTTTGCAGACACAGCTGGCTGTGGAGAG	2381
Q	y	1082	AlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAsp	1101
D	b	2382	GCCATTAAATCCCACTGGTTCTGCTCATTTTGGGAAGGAACAGAGGCCCATCTGCTGAT	2441
Q	y	1102	AspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGly	1121
D	b	2442	GAGATGAATTCGAATGGAAAGAAATCCCGCATTTTGGCAGTGGCATTCACACGGCTGGGG	2501



Db 202 CAGCTGGGGTGGAGTGGCCATTGGTTTCCTGGAGGGGCTTATTGGCCAGGACTT 261  
Qy 105 ThrArgHisGlyLysIleTrpLeuAspValSerCysTyrGlyAsnGluSerAlaLeu 124  
Db 262 -----GGCCCAATTTGGCTTTTGTATATCTTCATGTGAGGAGCAGAGTCAACTGTC 312  
Qy 125 TrpGluCysGlnHis-----ArgGluTrpGlySerHisAsnCysTyrHisGlyGlu 141  
Db 313 AGTGACTGTGAGCATTTCTAATATTAAAGACTATCGTAATGATGCTATAATCATGGTCGG 372  
Qy 142 AspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspGlyAsn 161  
Db 373 GATGCTGGAGTAGTCTGCTCAGGA-----TTTGTGCGTCTGGCTGGAGGGGAT 420  
Qy 162 AsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 181  
Db 421 GGACCTGTCTCAGGCGCAGTAGAAGTGCATTCGGAGAAGCTTGATCCCAAGTCTGTAT 480  
Qy 182 AspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSerSer 201  
Db 481 GGGAACTTCACATCTGCACCTGGCCAGATCATCTGTGCAGAGTTGGTGTGGCAAGGCT 540  
Qy 202 PheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAspAsp 221  
Db 541 GTGTCTGTCTGGGACATGAGCTCTTCAGAGAGTCCAGTGCACAGGCTCGGGCTGAAGAG 600  
Qy 222 IleLeuCysGlnGlyAsnGluAlaLeuLeuTrpAsnCysArgHisArgGlyTrpGlyAsn 241  
Db 601 TTCAGGTGTGAGGGGGAGGAGCCCTGAGCTCTGGGTCTGCCAGGAGTCCCTGTCCAGGG 660  
Qy 242 HisAspCysSerHisAsnGluAspValThrLeuThrCystyrAspSerSerAspLeuGlu 261  
Db 661 GGCACGTGTACACACAGTGGATCTGCTCAGGTGTGTGTTTCAGCATACTCAGAAGTCCGG 720  
Qy 262 LeuArgValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGly 281  
Db 721 CTC---ATGCAAAACGGCTCTCTCAGTGTGAAGGGAGGTGGAGATGAACATTTCTGGA 777  
Qy 282 ArgTrpGlyThrValCysHisLysTrpAsnAsnAlaAlaAspValValCysLys 301  
Db 778 CAATGGAGAGCGCTGTGTGCTCCCTGAGTGTGCCAATGCCAATGTTATCTGCTGT 837  
Qy 302 GlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySer 321  
Db 838 CAGCTCGCTGTGAGTTGCCATCTCCACCCCGGAGGACACACTTGGTGGAGAAGAGT 897  
Qy 322 AspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys 341  
Db 898 GATCAGATCTCTAACAGCCGATTTTCATGCTGTGGGCTGAGTCTCTCTGAGTGTGT 957  
Qy 342 ArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCys 361  
Db 958 CCTGTGACTGCCCTGGGTGGTCTGACTGTGCCATGGCAACACAGCTCTGTGATCTGC 1017  
Qy 362 SerAspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySerAsnAsnCysSerGlyArg 381  
Db 1018 TCAGGAAC----- 1026  
Qy 382 ValGluValArgIleHisGluGlnTrpTrpThrIleCysAspGlnAsnTrpLysAsnGlu 401  
Db 1027 ---CAGATCCAGGTGCTTCCCCAG----- 1047  
Qy 402 GlnAlaLeuValValCysLysGlnLeuGlyCysProPheSerValPheGlySerArgArg 421  
Db 1048 -----TGCACCACTCCGCTG----- 1062  
Qy 422 AlaLysProSerAsnGluAlaArgAspIleTrpIleAsnSerIleSerCysThrGlyAsn 441  
Db 1063 TCTCAACCTACAGGCTCTGCGGCGCTCA----- 1089  
Qy 442 GluSerAlaLeuTrpAspCysThrTyrAspGlyLysAlaLysArgThrCysPheArgArg 461

Db 1089 ----- 1089  
Qy 462 SerAspAlaGlyValIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAla 481  
Db 1090 GAGACAGGCCCCCTACTACTGCTCAGACAGCAGG-----CAGCTCCGCTGGTGACGGG 1143  
Qy 482 HisSerProCysTyrGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCys 501  
Db 1144 GCGGTCCTCGCGCGGAGAGTGGAGATCTTGACCAAGGCTCTCTGGGGCAGCATCTGT 1203  
Qy 502 HisAspArgTrpSerThrArgAsnAlaAlaValValCysLysGlnLeuGlyCysGlyLys 521  
Db 1204 GATCAGCGCTGGACCTGGACGATGCCCGCTGTGTGTGCAGGCAGCTGGGTGTGGAGAA 1263  
Qy 522 ProMetHisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIleTrpLeuAsp 541  
Db 1264 GCCCTCAATGCCAGGGGTCTGCTCACTTCGGGGCAGGATCAGGCCCATCTGGTTGGAC 1323  
Qy 542 AspValSerCysIleGlyAsnGluSerAsnIleTrpAspCysGluHisSerGlyTrpGly 561  
Db 1324 AACTTGAACTCACAGGAAGAGTCCACAGCTGTGGAGTGCCTTCCCGGGCTGGGG 1383  
Qy 562 LysHisAsnCysValHisArgGluAspValIleValThrCysSerGlyAspAlaThrTrp 581  
Db 1384 CAGCACAACTGCACACACAGCAGCGGGGTCTACTGTCTCA-----GAGTTCCTG 1437  
Qy 582 GlyLeuArgLeuValGlyLysSerAsnArgCysSerGlyArgLeuGluValTrpPheGln 601  
Db 1438 GCCCTCAGGATGGTGTGAGTGGAGCAGCAGCTGCTGGGTGGTGGAAATTTTCTACAA 1497  
Qy 602 GlyArgTrpGlyThrValCysAspGlyTrpAsnSerLysAlaAlaValValCys 621  
Db 1498 GGGACCTGGGCGACGTCTGCCGTAAACCCATGGAAGACATCACTGTCTCCACGATCTGC 1557  
Qy 622 SerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThr 640  
Db 1558 AGACAGCTGTGCTGTGGGACAGT-----GGAACCCCACTCTCTGTCTGTCT 1605  
Qy 641 -----GlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAspGluSer 657  
Db 1606 CTTAGAGAGGTTTATAGCCACAGTGGGTGATAGATCCAGTGTCCGAAAACACTGCACAC 1665  
Qy 658 AspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAsp 677  
Db 1666 TCTCTCTGGCAGTCTCTCTCTGACCTTGGAAATTACAACATCATGCTCTCCAAGAGGAA 1725  
Qy 678 ValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlySerSer 697  
Db 1726 GCCTATATCTGTGTGCAGACAGC-----AGACAGATCCCGCTGGTGGAGGTGGT 1779  
Qy 698 ArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsn 717  
Db 1780 CGCTGCTCTGGAGAGTGGAGATCTTCACACAGGCTCTCTGGGACCATCTGTGTATGAC 1839  
Qy 718 GlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGlyCysGlySerAlaIle 737  
Db 1840 CGCTGGGACCTGGAGTGCCTGCTGTGTGTGCAAGCAGCTGGGTGTGGAGAAAGCCCTG 1899  
Qy 738 ArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSer 757  
Db 1900 GACCCCATCTGTCTCTCTCTCTCGGACGGGATCAGGGCCCATCTGGCTGGATGAAGTG 1959  
Qy 758 GlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThr 777  
Db 1960 AACTGCAGAGGAGGAGTCCCAAGTATGGAGGTGCCCTTCTCTGGGATGGCGGCAACAC 2019  
Qy 778 AlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeu 797  
Db 2020 AACTGCAATCATCAAGAAGATGCAGGAGTCACTGCTCAGGATTT-----GTGCGTCTG 2073  
Qy 798 ValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArg 817  
Db 2074 GCTGGAGGATGACCCCTGCTCAGGCGGAGTAGAAGTGCAATTTCTGGAGAGACCTGGACC 2133

QY 818 SerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGluLeuAsn 837  
DB 2134 CCAGTGTCTGATGGAACTTACACTCCCACTCCAGTCCAGAGTGGGA 2193  
QY 838 CysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGlyValLeuThr 857  
DB 2194 TGTGGCAAGCTGTGCTCTCTGGACACATGCCATCCAGAGTCCAGAGGTC 2253  
QY 858 TrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877  
DB 2254 TGGGCTGAAGATTGAGTGTGATGGGGGAGCTGAGCTTGTCTGCCAGAGTG 2313  
QY 878 GlnHisProGluAspThrCysIleHisSerArgGluValGlyValValCysSerArgThr 897  
DB 2314 CCTGTCCAGGACATGTCTCCACAGTGGAGTGTCTGTTCAAGTGTAC 2373  
QY 898 ThrAspValArgLeuVal---AsnGlyLysSerGlnCysAspGlyGlnValGluLeuAsn 916  
DB 2374 ACAGAGTCCAGCTATGAAACACGACCTCTCAATGTGAGGGGAGGTGGAGATGAAG 2433  
QY 917 ValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgVal 936  
DB 2434 ATCTCTGGAGTGGAGACGCTCTGTGCTCCCACTGGAGTCTGCCAATGCCAATGT 2493  
QY 937 LeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysThrIleGly 956  
DB 2494 GTCTGTGTCAGTCTGGCTGTGGAGTGGCCATCTCCACCCAGAGGACACACTTGGTG 2553  
QY 957 GluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu 976  
DB 2554 GRAGGAGTGTACAGATCTACACAGCCCAATTTCACTGTCTAGGGCTGAGCTTCCG 2613  
QY 977 AspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrValSer 996  
DB 2614 TGGAGTGTCTGTGACTGCTTGGGTGGCTGAGTGTCCCATGGCAACACAGCTCT 2673  
QY 997 ValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAsp 1016  
DB 2674 GTGATCTGTCAAGAAACACACACCCAGGTGCTGCCACAGTGAACACTTCTGTCTCAA 2733  
QY 1017 ProTyrlSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeu 1036  
DB 2734 CTGCGAGCTCTGCGCCCTCAGAGGAGTGTCTCCCTACTGTCTACACAGCAGGAGCTC 2793  
QY 1037 ArgLeuValAspGlyAspSerArgCysAlaGlyArgValIleuIleThrHisAspGlyPhe 1056  
DB 2794 CGCTGTGGAGGGGGCGTCCCTCGCGGGGAGAGTGGAGATCTTGACACAGGCTCC 2853  
QY 1057 TrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLys 1076  
DB 2854 TGGGGCACCATCTGTGATGATGACTGGGACCTGGAGCATGCCGTGTGTGTGCGAGCAG 2913  
QY 1077 LeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGly 1096  
DB 2914 CTGGGTGTGGAGAGCCCTCAATGCCAGGGGTCTGCTCACTTCGGGGCAGATCAGGG 2973  
QY 1097 ProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysPro 1116  
DB 2974 CCTATCTGGCTGGACACCTGAATGCCACAGGAAAGAGTCCACGTGTGGAGGTGCCCT 3033  
QY 1117 SerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer 1136  
DB 3034 TCCCGGGGTGGGGCGGCGACACTCCAGACACAGAGGAGCGCGGGGTCTCTGCTCA 3093  
QY 1137 GluPheThrAlaLeuArgLeuThrSerGluThrGluThrGluSerCysAlaGlyArgLeu 1156  
DB 3094 GAGTCTCTGCCCTCAGGATGTGGAGC-----GAGGACACAGAGTGTCTGGGTGGCTG 3147  
QY 1157 GluValPheThrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIle 1176  
DB 3148 GAGTTTCTACAAACGGGACCTGGGCGAGTGTCTGCCGAGCCCATGGAGATATCACT 3207

QY 1177 AlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaPro 1196  
DB 3208 GTGTCGCGTATCTGCAGACAGTGTGGAGTGTGGAGACAGTGAAGTCTCAACACCTCTGT 3267  
QY 1197 LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHis 1216  
DB 3268 GTCTTCAGGGAAGTCTTAGACCCCGGTGGTAGATTTAATTCAGTGTGCGAANAATGGAT 3327  
QY 1217 IleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGlu 1236  
DB 3328 ACCTCTCTCTGCAGTGTCTTCTGCCCCATGGAATAACAGTTCATGCTCTCCAAAGGAG 3387  
QY 1237 GluThrTrpIleThrCysGlu----- 1243  
DB 3388 GAAGCTACATCTCATGTGAAGAAAGAACCCAGAGCTGCCAACTCTCCGCCCTGC 3447  
QY 1244 -----AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGlu 1260  
DB 3448 ACAGACAGACAGAGCTCCCTCAGGGAGGAGAGCAGGAGTCTCAGGGCGGTGGAG 3507  
QY 1261 IleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAla 1280  
DB 3508 GTGTGGCACAAAGCTCTCTGGGCGACCGTGTGCGATGACTCTCTGGAGCCTGGCAGAGCT 3567  
QY 1281 GluValValCysGlnClnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300  
DB 3568 GAGGTGTGTGTCAGCAGCTGGGCTGTGGCCAGGCTGGAAGCCTGGCGGTCTGCAGCA 3527  
QY 1301 PheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320  
DB 3628 TTTGGCCCTGGAATGGAGCATCTGGCTGGAGAGTGCAGTGGCGGGCGGGAGTCC 3687  
QY 1321 PheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAsp 1340  
DB 3688 TCCTCTGGGAGTGTCTTGGGAGCCTCTGGGCGAGGAGTGTGCAAGCAGCAGGAGGAT 3747  
QY 1341 AlaGlyValArgCysSerGly----- 1347  
DB 3748 GTGGGTGTGAGGTGCTGTGTAGGACAACTTCCCCAGCAGCAGCAGGAGCAGCA 3807  
QY 1348 -----GlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuLeu 1364  
DB 3808 ACAACCTCAAAATCTCTCCCTGGCATCTTCTCCCTGGGTCTCTGCTTATCTCTG 3867  
QY 1365 SerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPheLeuThrTrpCysArg 1384  
DB 3868 GGTCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3924  
QY 1385 ValGlnLysGlnLysHisLeuProLeuArgValSerThrArgArgArgGlySerLeuGlu 1404  
DB 3925 GCAGAGCGC-----AGAGCCTTATCCAGCTATGAAGATGCTCTGTCT 3966  
QY 1405 GluAsnLeuPheHisGluMetGluThrCysLeuLysArgGluAspProHisGlyThrArg 1424  
DB 3967 GRAGCTGTGTAGCAGAGCTCGATTACCTTCTGACACAGAGGAAGGTCTGGCGACCCCA 4026  
QY 1425 ThrSer---AspAspThrProAsnHisGlyCysGluAspAlaSerAspThrSerLeuLeu 1443  
DB 4027 GATCAGATGACTGATCTCCCTGATGAAATATGATGATGATGATGATGATGATGATGATG 4086  
QY 1444 GlyValLeuProAsnSerGluAlaThrLys 1453  
DB 4087 GGAACCTCTCTCTCCCTCTCAGGGGAATGAG 4116  
RESULT 12  
AAV49652  
ID AAV49652 standard; DNA; 5802 BP.  
XX  
AC AAV49652;  
XX  
DT 23-OCT-1998 (first entry)  
XX Human SRCR protein DNA.

XX  
KW Scavenger receptor cysteine rich domain; SRCR; diagnosis; treatment;  
KW nervous system; medullo-blastoma; glioma; breast; detection;  
KW autoantibody; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 107..5464  
FT /tag= a  
FT /product= SRCR protein"  
FT /note= "Scavenger receptor cysteine-rich domain  
FT containing protein"  
XX  
PN WO9830687-A2.  
XX  
PD 16-JUL-1998.  
XX  
XX 09-JAN-1998; 98WO-DE00096.  
XX  
PR 18-JUL-1997; 97DE-1030997.  
PR 09-JAN-1997; 97DE-1000519.  
XX  
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX  
XX Mollenhauer J., Poustka A;  
XX  
XX WPI: 1998-399136/34.  
DR P-PSDB; AAW64591.  
XX  
PT Proteins containing scavenger receptor, cysteine rich domain -  
PT useful for diagnosis and treatment of tumours  
XX  
PS Claim 4; Fig 2; 54pp; German.  
XX  
CC This sequence encodes a human protein which contains a SRCR (scavenger  
CC receptor, cysteine-rich) domain. The gene and encoded protein can be used  
CC to diagnose or treat tumours, particularly of the nervous system  
CC (medullo-blastoma or glioma) or breast. The DNA:sequence and probes  
CC derived from it, are used to identify genes that express SRCR-domain  
CC containing proteins, to determine the form in which these proteins exist  
CC and to assess the significance of individual forms on cellular  
CC properties. The protein can be used to detect the presence of  
CC autoantibodies and antibodies which regulate its expression.  
XX  
SQ Sequence 5802 BP; 1253 A; 1614 C; 1576 G; 1359 T; 0 other;  
Alignment Scores:  
Pred. No.: 1.37e-146 Length: 5802  
Score: 2435.00 Matches: 533  
Percent Similarity: 46.58% Conservative: 155  
Best Local Similarity: 36.09% Mismatches: 415  
Query Match: 29.92% Indels: 374  
DB: 19 Gaps: 28  
US-09-759-130b-381 (1-1453) x AAV49652 (1-5802)  
Qy 43 GlyThrAsp-----LeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60  
Db 389 GGATCTGATCTGTTGGCCCTCAGGCTGGTGAATGGAGATGGCAGGTGTGACGGCCGA 448  
Qy 61 ValGluValLysPheGlnGlnTrpGlyThrValCysAspGlyTrpAsnThrThr 80  
Db 449 GTGAGATCTATACCGAGGCTCTGGGGACCGCTGTGTGATGACAGCTGGGACACCAAT 508  
Qy 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMet----- 97  
Db 509 GATGCCAACGGTGGTCTAGGACAGCTGGGTGTGGTGGCCATGTCAGCTCCAGGAAT 568  
Qy 98 PheArgPheGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCys 117  
Db 569 GCCTGGTTGGCCAGGCG-----TCAGGACCCCATTCCTGGGATGATGTCGCGTGC 619

Qy 118 TyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCys 137  
Db 620 TCAGGACACGAATCTACCTGTGGAGCTGCCCCACAATGGCTGGCTCTCCCAATACTGT 679  
Qy 138 TyrHisGlyGluAspValGlyValAsnCysTyr----- 148  
Db 680 GGCCATGGTGAAGATGCTGGTGTATCTCTCAGCTGCCAGCCTCAGTCAACACTCAGG 739  
Qy 149 -----GlyGluAlaAsn 152  
Db 740 CCAGAAAGTTGGCCCTGTCCAGGATATCACCACCTGTACCACAGAAAGATCTGAATCCAGT 799  
Qy 153 LeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArgValGluValLysPhe 172  
Db 800 TTGGCCCTCAGGCTGGTGAATGGAGGCGACAGGTTGTCGAGGCCGAGTGGAGTCTTATAC 859  
Qy 173 GlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaVal 192  
Db 860 CGAGGCTCCTGGGCGACCGTGTGTGATGACTTACTGGGACACCAATGATGCCAATGTGTC 919  
Qy 193 CysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAla 212  
Db 920 TGCAGGCGAGCTGGGCTGTGGCTGGCCATGTTCAGCCCGAGGAATGCCAGTTGGCCAG 979  
Qy 213 ValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrp 232  
Db 980 GGCTCAGGACCCATGTCTGGATGTGGCTGCTCAGGACACGAGTCTTACCTGTGG 1039  
Qy 233 AsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeu 252  
Db 1040 AGCTGCCCGCCACAATGGCTGGCTCACCACCAACTGTGGCCATAGTAGAGAGCGTGGTGC 1099  
Qy 253 ThrCysTyr----- 255  
Db 1100 ATGCTCAGCTCCCGAGTCCCGCGCGACACCCAGCCAGATACTTGGCCCGACCTCAGAT 1159  
Qy 256 -----AspSerSerAspLeuGluLeuArgLeuValGlyGlyThrAsnArg 270  
Db 1160 GCATCAACAGCAGGACACCTGAATCCAGTTTGGCCCTCAGGCTGGTGAATGGAGGTGACAGG 1219  
Qy 271 CysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisLys 290  
Db 1220 TGTCCAGGCGGAGTGGAGGCTCTATACCGAGGCTCTGGGCGCAGCTGGCTGGCCACGCTCA 1279  
Qy 291 TrpAsnAlaAlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHis 310  
Db 1280 TGGACACACAGTGAGCGCAATGTGTCTCGCGGCGAGCTGGCTGGCCCGACGCTCA 1339  
Qy 311 PheAlaGlyLeuProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSer 330  
Db 1340 GCCCCAGGAATGCCCGGTTTGGCGCAGGGTTTCAGGACCCCAATGCTCGATGACGTGGCG 1399  
Qy 331 CysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnPheAsp 350  
Db 1400 TGCTCAGGCTATGAGTCTCTACCTGTGGAGCTGCCCGCCACAATGGCTGGCTGCCCATAAC 1459  
Qy 351 CysLeuHisGlnAsnAspValSerValIleCysSerAspGlyAlaAspLeuGluLeuArg 370  
Db 1460 TGTCCAGCAGCTGAAGACGCTGGTGTCTATCTGCTCAGCTGCC----- 1501  
Qy 371 LeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrp 390  
Db 1502 -----CACTCTCTGG 1510  
Qy 391 TrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeu 410  
Db 1511 TCGACG----- 1516  
Qy 411 GlyCysProPheSerValPheGlySerArgAlaLysProSerAsnGluAlaArgAsp 430  
Db 1517 -----CCGAGTCCAGACACATGTCGCG 1537  
Qy 431 IleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyr 450

Db 1538 ACCATCACTTCGCCGTCACAGTACAGTGTGAATCCAGT----- 1579  
Qy 451 AspGlyLysAlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSerAsp 470  
Db 1579 ----- 1579  
Qy 471 LysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGlu 490  
Db 1580 ----- 1580  
Qy 491 ValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsnAla 510  
Db 1631 GTCCATACCAAGGTCCTGGGGCACCCTGGTGCATGACAGCTGGGACACCAATGATGCC 1690  
Qy 511 AlaValValCysLysGlnLeuLeuCysGlyLysProMetHisValPheGlyMetThrTyr 530  
Db 1691 AATGTCGTCGACGACCCGGCTGGCTGGCCATGTCAGCCCGAGGAATCCCGG 1750  
Qy 531 PheLysGluAlaSerGlyProIleTrpLeuAspValSerCysIleCysAsnGluSer 550  
Db 1751 TTTGGTCAGGGCTCAGGACCCATTCCTGGATGATGTGCGCTGCTCAGGACACAGTCT 1810  
Qy 551 AsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAsp 570  
Db 1811 TACCCGCTGGAGCTGCCCCACCAATGGCTGCTCCCACTGTCGACATAGTGAAGAC 1870  
Qy 571 ValIleValThrCysSerGlyAspAla-----ThrTrp--- 581  
Db 1871 GCTGGTGTCTATCTGCTCAGCTCCAGTCCCGGCCCAACACCTAGTCCAGACACTGGCCA 1930  
Qy 582 -----GlyLeuArgLeuValGlyGly 588  
Db 1931 ACCTCACATGATCAACAGCAGGATCTGAATCCAGTTTGGCCCTGAGGCTGTGAATGA 1990  
Qy 589 SerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgTrpGlyThrValCys 608  
Db 1991 GTGACAGGTGTACGGGCGAGTGAGGTCTTATACCGAGGCTCCTGGGCGCACCGTGT 2050  
Qy 609 AspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysProSer 628  
Db 2051 GATGACTACTGGGACCAACCAATGATGCCAATGTGGTTTCAGGACGCTGGCTGTGCTGG 2110  
Qy 629 SerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAsp 648  
Db 2111 GCATGTCTACGCCAGCCAGAAATGCCGGTTTGGCCAGGTTTCAGGACCCATTCCTGGAT 2170  
Qy 649 AspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrpGly 668  
Db 2171 GATGTGCGTCTCAGGACATGAGTCTTATCTGTGAGCTGCCCCCAACAAATGGCTGGCTC 2230  
Qy 669 AsnAsnAspCysSerHisSerGluAspValGlyValIleCysSer----- 683  
Db 2231 TCCCAACACTGTGGCCATCATCAAGACGCTGTGTCTGCTCAGCTTCCAGTCCCGAG 2290  
Qy 684 -----AspAlaSer 686  
Db 2291 CCGACACCCAGCCACAGACTTGGCCAACTTGGCCATCATCAACAGCAGGATCTGAATCC 2350  
Qy 687 AspMetGluLeuArgLeuValGlyGlySerArgCysAlaGlyLysValGluValAsn 706  
Db 2351 AGTTTGGCCCTCAGGCTGGTGAATGGAGGTGACAGGTGTCAGGGCGGAGTGGAGGTCTTA 2410  
Qy 707 ValGlnGlyAlaValIleLeuCysAlaAsnGlyTrpGlyMetAsnIleAlaGluVal 726  
Db 2411 TACCGAGGCTCTGGGGCACCCTGTGTGATGACTACTGGGACCAACCAATGATCCCAATGTG 2470  
Qy 727 ValCysArgGlnLeuGluCysGlySerAlaIleArgValSerArgGluProHisPheThr 746  
Db 2471 GTTTCGAGCAGCTGGGCTGGCTGGGCGCCACGTCAGGCCCAAGAAATGCCCGGTTTGGC 2530  
Qy 747 GluArgThrLeuHisIleLeuMetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeu 766  
Db ----- 766

Db 2531 CAGGTTTCAGGACCCCATTTGCTGATGATGTGCGCTGCTCAGGACATGAGTCTATCTG 2590  
Qy 767 TrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAsnMetGluAlaSer 786  
Db 2591 TGGAGCTGCCCCCAATAGTGGTGTCTCCACAACTGTGGCCATCATCAAGACGCTGT 2650  
Qy 787 LeuIleCysSerAlaHisArg---GlnProArgLeuValGlyAlaAspMetProCysSer 805  
Db 2651 GTCATCTCTCAGCTTCCAGTCCAGTCCAGCGG----- 2680  
Qy 806 GlyArgValGluValLysHisAlaAspThrTrpArgSerValCysAspSerAspPheSer 825  
Db 2681 -----ACACCCAGCCACACTGTGGCCACG----- 2707  
Qy 826 LeuHisAlaAlaAsnValLeuCysArgGluLeuAsnCysGlyAspAlaIleSerLeuSer 845  
Db 2707 ----- 2707  
Qy 846 ValGlyAspHisPheGlyLysGlyAsnGlyLeuThrTrpAlaGluLysPheGlnCysGlu 865  
Db 2708 -----TCTGTGTCATCAACAGCA 2725  
Qy 866 GlySerGluThrHisLeuAlaLeuCysProIleValGlnHisProGluAspThrCysIle 885  
Db 2726 GGATCTGAATCCACTTTGGCCCTG----- 2749  
Qy 886 HisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValAsnGly 905  
Db 2750 -----AGACTGTGAATGA 2764  
Qy 906 LysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCys 925  
Db 2765 GGTGACAGGTGTGAGGCGGAGTGGAGTCTATACCAAGGCTCCTGGGCGCACCGTGT 2824  
Qy 926 AspThrHisTrpAspProGluAspAlaArgValLeuCysArgGlnLeuSerCysGlyThr 945  
Db 2825 GATGACTACTGGGACCAACCAATGATGCCAAGTGGTGTGACGACACTGGGCTGTGGCTG 2884  
Qy 946 AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHis 965  
Db 2885 GCATGTGACGCCCGCAGAAATGCCAGTTTGGCCAGGCTCAGGACCCATTCCTGGAT 2944  
Qy 966 ArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGly 985  
Db 2945 GATGTGCGCTGTCTAGGACACAGTCTTACCTGTGGAGCTGCCCCACAAATGGCTGGCTC 3004  
Qy 986 AlaProProCysIleHisGlyAsnThrValSerValIleCysThrGlySerLeuThrGln 1005  
Db 3005 TCCCAACACTGTGGCCATCATGAAGTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 3064  
Qy 1006 -----ProLeuPheProCysLeuAlaAsnValSerAspProTyrLeuSerAlaValPro 1023  
Db 3065 TCAAGCCCGCAGCCAGACTACTTGGCTGACCACTTACCGGCATTCAGCATAGATCT 3124  
Qy 1024 GluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSer 1043  
Db 3125 GAATCCAGTTTG-----GCTGTAGGCTGTGAATGGAGGTGAC 3163  
Qy 1044 ArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAsp 1063  
Db 3164 AGGTCTCAGGCGGAGTGGAGGTCTCTATCGAGGCTCTCTGGGGAACCGTGTGTGATGAC 3223  
Qy 1064 GlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPhe 1083  
Db 3224 AGCTGGGACACCAATGATGCCAATGTGTCTGACGAGCTGGGCTGGCTGGGCGCATG 3283  
Qy 1084 AsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspLeu 1103  
Db 3284 TCGGCGCCAGGAATGCGCGGTTTGGCCAGGCTCAGGACCCATTCCTGATGATGTG 3343  
Qy 1104 AsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHis 1123  
Db ----- 1123  
Db 3344 CGCTCTCAGGGAATGAGTCTTACCTGAGCTGCCCGCCCAAGGCTGGCTCAGCCAC 3403

Oy 1124 AspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeu 1143  
: : : : :  
Db 3404 AACTGTGCCCATCAGAGAGCGTGGTCTCATCTGCTCA-----GCCACCAATA 3454  
Oy 1144 TyrSerGluThr-----GluThrGlu 1150  
: : : : :  
Db 3455 AATCTACTACGACAGATGTTGGCATCCAAACAACACTACACCACTGCCAAGACCTCTTCA 3514  
Oy 1151 SerCysAlaGlyArgLeuValPheTyr---AsnGlyThrTrpGlySer- 1166  
: : : : :  
Db 3515 AATGTGTGGCTTCTTA-----TTCTATCCAGTGGGACATCTCTCCAGCCCATCTAC 3568  
Oy 1166 ----- 1166  
Db 3569 CTGCATACTACCCCAACAATGCTAAGTGTCTTTGGGAAATAGAAGTGAATCTGGTTAT 3628  
Oy 1167 -----ValGlyArgAsnIle----- 1172  
Db 3629 CGCATAAACCTGGGCTTCAGTGAATCTGAAATGGAGGCACACCAATACTGCAGTTTGTAT 3688  
Oy 1173 -----ThrThrAlaIleAlaGlyIleValCys--- 1181  
: : : : :  
Db 3689 TATGTTGAATCTTGTATGATCATTAATGATAGCAGTCTCTGCTGGGAAATCTGTAAT 3748  
Oy 1182 -----ArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeu-----AlaPro 1196  
: : : : :  
Db 3749 GATACGAGCAATATTTACATCTCTTACAAACCAATGACCATTCACATTTTGAAGTAC 3808  
Oy 1197 LeuSerLysThrGlySerGlyPheMet---TrpValAspAlaGlnCysProLysThr 1215  
: : : : :  
Db 3809 ATCAGTTTCCAAAACACTGGGCTTTTGGCTTGGTATAC----- 3847  
Oy 1216 HisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgIleSerSerProAla 1235  
: : : : :  
Db 3848 -----TCTTCCCAAGC 3859  
Oy 1236 GluGluThrTrpIleThrCysGluAspArgIleArgValArgGlyAspThrGluCys 1255  
: : : : :  
Db 3860 GATCCACCTTG-----AGGTTGCTCAATTTAAATTCATCTATGCTATGT 3907  
Oy 1256 SerGlyArgValGluIleThrPheAlaGlySerTrpGlyThrValCysAspAspSerTrp 1275  
: : : : :  
Db 3908 GCCGGCGGTGAGAAATTTACATGGTGGCCACCTGGGGAGAGTTGTGATGATCTCCG 3967  
Oy 1276 AspLeuAlaGluAlaGluValValCysGlnLeuGlyCysGlySerAlaLeuAlaAla 1295  
: : : : :  
Db 3968 ACCATTACGAGAGTGGTGTGTCAGACAGCTAGGCTGGACGTGCAGTTTCACCC 4027  
Oy 1296 LeuArgAspAlaSerPheGlyGlnGlyThrGlyThrIleTrpLeuAspMetArgCys 1315  
: : : : :  
Db 4028 CTTGGAATGATATTTTGGCTCTGGCTCTGGCTCTGGCCCATCACCCTGGACGATGACAGTGC 4087  
Oy 1316 LysGlyAsnGluSerPheLeuTrpAspCysHisAlaLysProTrpGlyLysSerAspCys 1335  
: : : : :  
Db 4088 TCAGGAGGAGAAATCCTCTCTGGCAGTGGCCGAGACCGAGCTGGTTCTCCCAACATGT 4147  
Oy 1336 GlyHisLysGluAspAlaGlyValArgCysSerGlyGlnSerLeuLysSer 1352  
: : : : :  
Db 4148 AATCATCTGAAGTGTGGTGTCTATCTGCTCAGGAACCATCATCTATCGACA 4198  
RESULT 13  
AAS17590  
ID AAS17590 standard; cDNA; 2697 BP.  
XX  
AC AAS17590;  
XX  
DT 26-FEB-2002 (first entry)  
XX  
DE DNA encoding novel secreted protein #19.  
XX  
KW Secreted protein; cytostatic; immunosuppressive; vulnery; vaccine;  
antlinflammatory; neuroprotective; nephrotropic; cardiovascular;

KW human; cancer; autoimmune disease; wound healing disorder; infection;  
KW haematopoietic disorder; inflammatory disorder; infertility;  
KW neurological disease; psychiatric disease; cardiovascular disease;  
KW respiratory disease; renal; gastrointestinal; ss.  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT 1..2697  
FT CDS /\*tag= a  
FT /product= "Human secreted protein"  
XX WO200179454-A1.  
PN  
XX 25-OCT-2001.  
PD  
XX 11-APR-2001; 2001WO-US11797.  
PF  
XX 13-APR-2000; 2000US-196603P.  
PR 24-APR-2000; 2000US-199417P.  
PR  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX (SMIK ) SMITHKLINE BEECHAM PLC.  
PA  
XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;  
PI WPI: 2002-061975/08.  
XX P-PSDB: AAU09878.  
DR  
XX New secreted proteins or polypeptides, useful for treating e.g. cancer,  
PT autoimmune diseases, wound healing disorder, infections, haematopoietic  
PT disorders, inflammatory disorders, infertility, cancer.  
XX  
PS Claim 2; Page 47-48; 92pp; English.  
XX

CC The invention relates to an isolated novel secreted polypeptide (I) and  
CC polynucleotide (II). (I) and (II) are useful for treating cancer,  
CC autoimmune diseases, wound healing disorder, infections, haematopoietic  
CC disorders, inflammatory disorders, infertility, neurological and  
CC psychiatric diseases, cardiovascular diseases, respiratory diseases,  
CC renal diseases, or gastrointestinal diseases. These may also be used to  
CC treat diseases, abnormalities and disorders caused by abnormal  
CC expression, production, function and/or metabolism of the genes, as  
CC vaccines for inducing immunological response in a mammal, and in  
CC screening methods for detecting the effect of added compounds on the  
CC production of mRNA and polypeptide in cells. The polypeptides can be used  
CC as immunogens to produce antibodies immunospecific for the polypeptides,  
CC and to identify membrane-bound or soluble receptors. The polynucleotides  
CC may be used as diagnostic reagents, in chromosome localisation studies,  
CC and in tissue expression studies. The present sequence represents the  
CC coding sequence of novel human secreted protein #19.  
XX

SQ Sequence 2697 BP; 358 A; 841 C; 1055 G; 443 T; 0 other;

Alignment Scores:  
Pred. NO.: 6.94e-104 Length: 2697  
Score: 1763.50 Matches: 361  
Percent Similarity: 49.37% Conservative: 108  
Best Local Similarity: 38.00% Mismatches: 262  
Query Match: 21.67% Indels: 219  
Gaps: 16

US-09-759-130B-381 (1-1453) x AAS17590 (1-2697)

Oy 536 GlyProIleTrpLeuAspValSerCysIleGlyAsnGluSerAsnIleTrpAspCys 555  
Db 31 GGGCCCCCTCTCTG-----AATCTCTGGGCAGTC 60  
Oy 556 GluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAspValIleValThrCys 575  
Db 61 CCCATTGGTGACACAGT----- 78  
Oy 576 SerGlyAspAlaThrTrpGlyLeuArgLeuValGlySerAsnArgCysSerGlyArg 595

Db	79	-----GCTGTGAGGCTGGCGTACAGACACAGCAGCTGCGACGAGTG	120
Qy	596	LeuGluValTyrPheInclnYArgTrpGlyThrValCysAspGlyTrpAsnSerLys	615
Db	121	GTGTGGTCCGACACACGGGCGCATGGGATACGTGTGTCAACAGAGTAGTGACGTGGCA	180
Qy	616	AlaAlaAlaValValCysSerGlnLeuAspCysProSerSerIleIleGlyMetGlyLeu	635
Db	181	GAGCCCTCTGCTGTGCAGGACGTGGGCTGCGGCCCTGCCGTGGCGGCCCCCAAGTAT	240
Qy	636	GlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAspValSerCysAspGlyAsp	655
Db	241	GTCCCGCTGCTGGAGAGATGCCCGCAGCCCTGGCTTCAACAGCTGTCTGCCGGGGCAAC	300
Qy	656	GluSerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSer	675
Db	301	GAGTCCTCCTCTGGGAGGTGACCCCTTGGCTCATGGTGCCAGACGCCGTGCCCCACGCA	360
Qy	676	GluAspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGly	695
Db	361	TGGGTGGTGGTCCGCTGTGCTCCAACCGCATTTCCGGGAGCTTCCGGCTGGTGAAGGC	420
Qy	696	SerSerArgCysAlaGlyLysValGluAsnValGlnGlyAlaValGlyIleLeuCys	715
Db	421	CGCAGTCCCTCGCGGACATCCCGGAGATCAGAACAGTGAATGGGTGGAGCCGCTCTGT	480
Qy	716	AlaAsnGlyTrpGlyMetAsnIleAlaGlu-----ValValCysArgGlnLeuGluCys	733
Db	481	GTC-----CTGCATGTGAGGAGGCCATGGTGTCTCCCGGGAGTGGGGTGC	528
Qy	734	GlySerAlaIleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeu	753
Db	529	GGCCCTGTCTCAGGCCCCCGC-----CGGACGTGGGCGTGC	570
Qy	754	MetSerAsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGlu	773
Db	571	AGGAAGTACCTGGCTTCGAGGGGTACCGAGGCCACCATCCCGACGCTCAGACATGGAAC	630
Qy	774	TrpIysGlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArg	793
Db	631	AACTTCGCGACGGCTGGACCTGGCGCTGGACGAGGTGGTCTGCTCAGACACACC	690
Qy	794	GlnProArgLeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAla	813
Db	691	GAGGCCGAGTGTGGCGCGGAGCACCCCTCGCGCGGCGCCCTGGAGTG-----	741
Qy	814	AspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCys	833
Db	742	---ACCTGGGGCACCGTCTGTATCGGCCCTGGACCTGGCCACCGGCCACCGTGGTGC	798
Qy	834	ArgGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGly	853
Db	799	CGGGAGCTGAGTGTGGGCGGTGTGTGTCCAGCCCGAGGGCGCCCGCTTCGGCGGGGC	858
Qy	854	AsnGlyLeuThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeu	873
Db	859	TCGGGCGGTGTGGACGAGGCCCTTCGCTGTGGGCGCAACGAGTCCGCTGCTTCCAC	918
Qy	874	CysProIleValGlnHisProGluAspThrCysIleHisSerArgGluValGlyValVal	893
Db	919	TGCCCA-----CGGGGGGTGGAGCCAGTGTGGGCATGTCACGACGGGGCTCAGG	972
Qy	894	CysSerArgTyrThrAspValArgLeuValAsnGlyLysSerGlnCysAspGlyGlnVal	913
Db	973	TGCTCAGAGTT-----AGGATGGTCAACGGCAGCAGCAGTGTGAGGGCGCGGTG	1023
Qy	914	GluIleAsnValLeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAsp	933
Db	1024	GAGTTCAGGTGCGAGGGTCTGGGCACCCCTCTGTGGCCACCCTGGACATAGCAGAT	1083
Qy	934	AlaArgValLeuCysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLys	953

Db	1084	GCACCGTCCTCTGCCACAGCTCAACTGTGGCAACCGCGGTGGCCGCCACCTGGAGGAGGC	1141	
Qy	954	TyrIleGlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGlu	973	
Db	1144	CATTTTGGGACGGGACGCTGCCATCTGGCCCTGATGCCTTTTCACTGTGAGGGACAGAG	1203	
Qy	974	SerLeuLeuAspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsn	993	
Db	1204	TCCTACTTGTGAATGTCCAGTAAGACACCTGGGGGCCCGCCCTGTGCCCGGGAAC	1263	
Qy	994	ThrValSerValIleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsn	1013	
Db	1264	ACAGCCTCCGCGTCTGCTCAGGT	1296	
Qy	1014	ValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAsp	1033	
Db	1297	GCC	1299	
Qy	1034	LysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIle	1051	
Db	1300	CTCGACTGAGGAAGGACAGACGCGTGTGACGGCGCGCTGGAGTCTCCCTG	1353	
Qy	1051		1051	
Db	1354	GATGCGTGTGGGGCGCGTCTCTGGACAGATGCTGGGACCTGCGCGCGCGCGGTGGTG	1413	
Qy	1051		1051	
Db	1414	TGCCGGCACTCGGTGTCAGAGGGGGCCAGCAGCCATATGACGCACCTGCCCCACGCCG	1473	
Qy	1051		1051	
Db	1474	GGATCGTCCAGTGGCGCTGAGCCGCTGCTGGCACCCGACCCGCCCTGACT	1533	
Qy	1051		1051	
Db	1534	CAGTGCACAGTGTCCGACCCCTGACAGGACCCCGCGGGGACCTGGCGGGACCGCGGCTG	1593	
Qy	1051		1051	
Db	1594	GTGTGCTCGGTGAGTTCGGAACCGCTGCCCATGGCCGTGCCACGGGATCCCGGCG	1653	
Qy	1051		1051	
Db	1654	GCCTGACTGTCTCTCCACAGGGAGCCTCAGGGTGGCGCTGGCGCGGGCGCGGGCGG	1713	
Qy	1052	TyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAlaHis	1071	
Db	1714	CTGCAGGGGGCGGTGGGACACCGTGTGACGATGCTTGGGACCTCGGGGACGGCAC	1773	
Qy	1072	ValValCysGlnIysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPhe	1091	
Db	1774	GTGGTCTGACGACGTGGCTGTGGCGCGCCCTGAGCGCCCTGGGGGCCGACATTC	1833	
Qy	1092	GlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHis	1111	
Db	1834	GGAGCCGGGCGACGGGCGCATCTGCCTGGACAGCTGGGCTGCCAGGGCCACGATCTCGC	1893	
Qy	1112	LeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAla	1131	
Db	1894	CTGTGCAGTGGCCGTCGGCGGCTGGGGCGGCGACGACCTGAGGACACAGGAGGACGCC	1953	
Qy	1132	GlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSer	1151	
Db	1954	GGCGTCTTCTGCTCAGATCGGTGGCTCTGAGGCTG	2007	
Qy	1152	CysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArgAsn	1171	
Db	2008	TGTGCTGGGTGGCTGACGCTTCTACATGAGGACCTGGGGCGCCATGTGCAGCAATGCC	2067	
Qy	1172	IleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal	1191	
Db	2068	CTGAAGGACCTCTCTTGTGCCATCATCTCTTCTCAACAGCAGCTGGGGTGTGGGGTGGGGA	2124	



Db 386 TGAGTCAGCTCTCTGGGAATGTCACACCGGGAATGGGAAGCCATACTGTTATCATGG 445  
 Qy 140 yGluAspValGlyValAsnCysTyrGlyGluAlaAsnLeuGlyLeuArgLeuValAspG1 160  
 Db 446 AGAAGAGTGTGTGTAAGTCTTACCGGTGAAGCCATCTGGGTTTGAGGCTAGTGGATGG 505  
 Qy 160 yAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTyrGlyThrIleCy 180  
 Db 506 AAACAC-TCCGTTCAGGAGAGTGGAGTGAATTCACAGAAAGTGGGGAACATATATG 564  
 Qy 180 sAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSe 200  
 Db 565 TGATGATGGGTGAATCTTAATACCCCTCCGCTCTGTCAGGCAACTAGGATGCCATC 624  
 Qy 200 rSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrpLeuAs 220  
 Db 625 TCTCTTTATCTCTGAGTGTCTAACAGCCCTGCTGTATGCGCCCATTTGGCTGGA 584  
 Qy 220 pAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpG1 240  
 Db 685 TGACATTTATGCCAGGGGAATGAGTGGCACTCTGGAATTCGACACATCTGGATGGG 744  
 Qy 240 yAsnHisAspCysSerHisAsnGluAspValThrLeuThrCysTyrAspSerSerAspLe 260  
 Db 745 AAATCATGACTGACGTACATGAGGATGTCACATTAATCTGTTATGATAGTAGTGTCT 804  
 Qy 260 uGluLeuArgLeu 264  
 Db 805 TGAACGTAGGCTC 817

RESULT 15  
 AAH99258  
 ID AAH99258 standard; cDNA; 690 BP.  
 AC AAH99258;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Human protein encoding cDNA sequence SEQ ID NO:93.  
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 KW antinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 KW antibacterial; endocrine; cardiant; central nervous system; virucide;  
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
 KW antiagregant; haemostatic; vulnery; antidiabetic; osteopathic; eczema;  
 KW dermatologic; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 KW immunostimulant; gene therapy; antisenese therapy; vaccine; inflammation;  
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune;  
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 KW neurological disorder; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200153455-A2.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 22-DEC-2000; 2000WO-US35017.  
 XX  
 PR 23-DEC-1999; 99US-0471275.  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX

DR

WPI; 2001-457603/49.

P-PSDB; AAH25317.

XX

Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

Claim 1; Page 341; 1217pp; English.

PS

XX

AAH99166 to AAH99904 encode the human proteins given in AAH25225 to AAH25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antinflammatory; antirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiagregant; haemostatic; vulnery; antidiabetic; cytostatic; neuroprotective; antiparkinsonian; antiasthmatic; antidiabetic; osteopathic; dermatologic; antiallergic; antiasthmatic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisenese therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders.

XX

SQ Sequence 690 BP; 163 A; 134 C; 207 G; 186 T; 0 other;

Alignment Scores:

Pred. No.:	2,42e-67	Length:	690
Score:	1183.00	Matches:	214
Percent Similarity:	99.07%	Conservative:	0
Best Local Similarity:	99.07%	Mismatches:	2
Query Match:	14.54%	Indels:	0
DB:	22	Gaps:	0

US-09-759-130B-381 (1-1453) x AAH99258 (1-690)

Qy

636 GlyAsnAlaSerThrGlyTyGlyLysIleTrpLeuAspValSerCysAspGlyAsp 655

Db

43 GAAACGCTCTTACAGGATATGGAATAATTTGGCTCGATGATGTTCTTCGTATGGAGAT 102

Qy

656 GluSerAspLeuTrpSerCysArgAsnSerGlyTrpGlyAsnAsnAspCysSerHisSer 675

Db

103 GAGTCAGATCTCTGCTCATGCAGGAACAGTGGTGGGAAATAATCACTGCAGTCACACT 162

Qy

676 GluAspValGlyValIleCysSerAspAlaSerAspMetGluLeuArgLeuValGlyGly 695

Db

163 GAAGATGTTGGAGTATCTGTTCTGATGTCGATATGGAGCTGAGGCTTGTGGTGGA 222

Qy

696 SerSerArgCysAlaGlyLysValGluValAsnValGlnGlyAlaValGlyIleLeuCys 715

Db

223 AGCAGAGGTGTGCTGGGAAATGAACATTGCTGAAGTTGTTTTCAGGCAACTTGAATGTGCT 282

Qy

716 AlaAsnGlyTrpGlyMetAsnIleAlaGluValValCysArgGlnLeuGluCysGlySer 735

Db

283 GCTAATGGCTGGGGAATGAACATTGCTGAAGTTGTTTTCAGGCAACTTGAATGTGCT 342

Qy

736 AlaIleArgValSerArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSer 755

Db

343 GCAATCAGGGTCTCCAGAGAGCCTCATTTACAGAAAGAACATTACACATTTAATGTCC 402

Qy

756 AsnSerGlyCysThrGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLys 775

Db

403 AATTCCTGGCTGCTGGAGGGAAGCCTCTCTCTGGATGTATGATGATGGAGTGGGAAA 462

Qy

776 GlnThrAlaCysHisLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnPro 795

Db

463 CAGACTGCGTGTCAITTAATAATATGGAAGCAAGTTTGATCTGCTCAGCCACAGCCGCC 522

Qy 796 ArgLeuValGlyAlaAspMetProCysSerGlyArgValGluValLysHisAlaAspThr 815  
|||||  
Db 523 AGGCTGGTTGGAGCTGATATGCCCTGCTCTGGACGTGTGAAGTGAACATGCACACACA 582  
|||||  
Qy 816 TrpArgSerValCysAspSerAspPheSerLeuHisAlaAlaAsnValLeuCysArgGlu 835  
|||||  
Db 583 TGGCGCTCTGTCTGTGATTCTCTCTTCATGCTGCCAATGTGCTGTGCAGAGAA 642  
|||||  
Qy 836 LeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGly 851  
|||||  
Db 643 TTAACCTGTGGAGATGCCAATATCTTCTCTGGGAGATCATTGGG 690  
|||||

Search completed: May 12, 2003, 06:58:24  
Job time : 670 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 2862 seconds  
(without alignments)  
8222.237 Million cell updates/sec

Title: US-09-759-130b-381  
Perfect score: 8138  
Sequence: 1 MMLPONSWHIDFGRCCHON.....CEDASDTSLGLVPASEATK 1453

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus.p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09759130/runat\_06052003\_123554\_23365/app\_query.fasta\_1.1607  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09759130.ecgn\_1.1.2168 @runat\_06052003\_123554\_23365 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*\*  
1: em\_estba:\*\*  
2: em\_esthum:\*\*  
3: em\_estlin:\*\*  
4: em\_estnu:\*\*  
5: em\_estov:\*\*  
6: em\_estpl:\*\*  
7: em\_estro:\*\*  
8: em\_estc:\*\*  
9: gb\_estti:\*\*  
10: gb\_est2:\*\*  
11: gb\_htc:\*\*  
12: gb\_est3:\*\*  
13: gb\_est4:\*\*  
14: gb\_est5:\*\*  
15: em\_estfun:\*\*  
16: em\_estom:\*\*  
17: gb\_gss:\*\*  
18: em\_gss\_hum:\*\*  
19: em\_gss\_inv:\*\*  
20: em\_gss\_pln:\*\*  
21: em\_gss\_vrt:\*\*  
22: em\_gss\_fun:\*\*  
23: em\_gss\_mam:\*\*  
24: em\_gss\_mus:\*\*  
25: em\_gss\_other:\*\*  
26: em\_gss\_pro:\*\*  
27: em\_gss\_rod:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1765.5	21.7	999	9	AL529439	AL529439 AL529439
2	1083	13.3	940	9	AL550994	AL550994 AL550994
3	1072	13.2	582	12	BG570972	BG570972 602591546
4	960	11.8	799	13	BI185454	BI185454 UNL-P-FN-
5	942.5	11.6	1108	9	AL544534	AL544534 AL544534
6	936.5	11.5	1100	9	AL540905	AL540905 AL540905
7	929	11.4	718	13	BI183796	BI183796 UNL-P-FN-
8	895	11.0	554	9	AL529438	AL529438 AL529438
9	836.5	10.3	930	13	BI838644	BI838644 603086185
10	767	9.4	1153	9	AL540904	AL540904 AL540904
11	739	9.1	723	13	BI836459	BI836459 603082833
12	718	8.8	644	13	BI009838	BI009838 MR4-RT004
13	701	8.6	585	14	BQ327934	BQ327934 MR4-RT004
14	695	8.5	653	10	BE502724	BE502724 hz78412.x
15	679	8.3	616	13	BI009795	BI009795 MR4-RT004
16	675	8.3	652	10	BB622216	BB622216 BB622216
17	669.5	8.2	913	14	BQ710454	BQ710454 AGENCOURT
18	667.5	8.2	711	14	BQ378694	BQ378694 MR3-UT009
19	667	8.2	574	14	BQ327932	BQ327932 MR4-RT004
20	666	8.2	581	14	BQ776070	BQ776070 UT-H-FHO-
21	660.5	8.1	593	14	BQ327876	BQ327876 MR4-RT004
22	658	8.1	574	13	BI009286	BI009286 MR4-RT004
23	657	8.1	583	14	BQ327909	BQ327909 MR4-RT004
24	651	8.0	596	13	BI009802	BI009802 MR4-RT004
25	649	8.0	934	14	BQ936228	BQ936228 AGENCOURT
26	645.5	7.9	927	9	AL570629	AL570629 AL570629
27	634.5	7.8	689	12	BF096204	BF096204 MR3-UT009
28	634	7.8	2925	11	AK014331	AK014331 Mus muscu
29	626	7.7	505	13	BI038644	BI038644 IL3-NT028
30	626	7.7	705	17	BI6308	BI6308 34707.TP.CI
31	625	7.7	607	14	BQ327933	BQ327933 MR4-RT004
32	624	7.7	3665	11	BC035470	BC035470 Homo sapi
33	621	7.6	523	12	BG334729	BG334729 602533892
34	620	7.6	569	10	AW631841	AW631841 90867.MAR
35	616.5	7.6	833	9	AL574811	AL574811 AL574811
36	616	7.6	560	13	BI009799	BI009799 MR4-RT004
37	612	7.5	529	13	BI009105	BI009105 MR4-RT004
38	612	7.5	531	14	BQ327873	BQ327873 MR4-RT004
39	612	7.5	579	14	BQ327906	BQ327906 MR4-RT004
40	607	7.5	602	12	BG641637	BG641637 PG11c.pko
41	605	7.4	583	14	BQ368412	BQ368412 PN3-GN051
42	587.5	7.2	594	10	AV616074	AV616074 AV616074
43	579	7.1	822	12	BF792807	BF792807 602253274
44	578	7.1	1581	11	BC020481	BC020481 Homo sapi
45	577.5	7.1	850	9	AL550308	AL550308 AL550308

ALIGNMENTS

RESULT 1  
AL529439  
LOCUS  
DEFINITION AL529439 LTI\_NFL001\_NBC4 Homo sapiens cdna clone CS0DD006Y005 5  
prime, mRNA linear EST 13-FEB-2001  
ACCESSION AL529439  
VERSION AL529439.1 GI:12792932  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 999)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization

JOURNAL  
COMMENT

Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source

Location/Qualifiers  
1. .999  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="LTI\_NFL001\_NBC4"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 220 a 236 c 318 g 224 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 5,88e-168 Length: 999  
Score: 1765.50 Matches: 327  
Percent Similarity: 98.50% Conservative: 1  
Best local Similarity: 98.20% Mismatches: 5  
Query Match: 21.69% Indels: 3  
DB: 9 Gaps: 0

US-09-759-130b-381 (1-1453) x AL529439 (1-999)

Qy 1060 ileCysAspArgpTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGlyCys 1079  
Db 1 ATCTGTGATGACGGCTGGGACCTGAGCGATGCCACGTGGTGTGTCAAAAGCTGGGCTGT 60  
Qy 1080 GlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluCysGlySerProIleTrp 1099  
Db 61 GGAGTGGCTTCAATGCCACGGTCTCTGCTCACTTTGGGGAGGGGTGAGGCCCATCTGG 120  
Qy 1100 LeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArgGly 1119  
Db 121 CTGGATGACCTGAACCTGCACAGGAATGGAGTCCCACTTGTGGCAGTGCCTTCCCGCGGC 180  
Qy 1120 TrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPheThr 1139  
Db 181 TGGGGCAGCAGCACTGCAGGCACAGGAGGACGCGAGGGGTCTATCTCTCAGAATTCACA 240  
Qy 1140 AlaLeuArgLeuTyTrSerGluThrGluThrGluSerCysAlaGlyValGluValPhe 1159  
Db 241 GCCTTGAGGCTCTACAGTGAACCTGAAACAGAGAGCTGTGCGGAGATGGAGTCTTC 300  
Qy 1160 TyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIle 1179  
Db 301 TATAACGGACCTGGGGCAGCGTGGCAGAGGAACATCACACAGCCATAGCAGGCATT 360  
Qy 1180 ValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSerLys 1199  
Db 361 GTGTGAGCAGCAGTGGGCTGTGGGGAATGGAGTAGTCAGCTCGCCCTTATCTAAG 420  
Qy 1200 ThrGlySerGlyPheMetTrpValAspAlleGlnCysProLysThrHisIleSerIle 1219  
Db 421 ACAGGCTCTGGTTTCATGTGGTGGATGACATTCAGTGTCTTAAACAGCATATCTCCATA 480  
Qy 1220 TrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThrTrp 1239  
Db 481 TGGCAGTGCCTGTCTGCCCCATGGGAGCGCAAGAATCTCCAGCCCGCAGCAGACCTGG 540

Qy 1240 IleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgVal 1259  
Db 541 ATCATGTGAAGATAGAAATAAGAGTGGTGGAGGAGACACCGAGTCTCTGGGAGAGTG 600  
Qy 1260 GluIleTrpHisAlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGlu 1279  
Db 601 GAGATCTGGCAGCGAGGCTCTCTGGGGCACAGTGTGTGATGACTCTGGGACCTGGCCGAG 660  
Qy 1280 AlaGluValValCysGlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAla 1299  
Db 661 GCGGAAGTGGTGTGCACAGCTGGCTGGCTCTCTGGCTGGCTGGCTGGAGGACGCT 720  
Qy 1300 SerPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGlu 1319  
Db 721 TCGTTTGGCCAGGGAACCTGGGAACCATCTGGTTGGATGACATCGCTGGCTGCAAGAAATGAG 780  
Qy 1320 SerPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGlu 1339  
Db 781 TCATTTCTATGGAGCTGTACCCCAACCCCTGGGACAGAGTGTGGACACCAAGNA 840  
Qy 1340 AspAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHis 1359  
Db 841 GATGCTGGCTGAGGTGCTCTGGACAGTCGCTGAAATCACTGAATGCCTCCTCAGGTCAAT 900  
Qy 1360 LeuAlaLeuLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIleLeuPhe 1379  
Db 901 TTAGCACTTATTTATCCAGTATCTTTGGGCT-CTTCTCTGGTCTGTATAT-CTATTT 958  
Qy 1380 LeuThrTrpCysArgValGlnLysGlnLysHisLeuPro 1392  
Db 959 CTCACGTGTGGCGAGTTCAGAA-WAAAAACATCTGCC 996  
RESULT 2  
AL550994 940 bp mRNA linear EST 16-FEB-2001  
LOCUS AL550994 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0DI066YE07 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL550994  
VERSION AL550994.1 GI:12888516  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1. (bases 1 to 940)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
Location/Qualifiers  
1. .940  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: pCMVSPORT 6; Site: 1; NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 228 a 212 c 279 g 218 t 3 others  
ORIGIN

Alignment Scores:

Pred. No.: 1,03e-98 Length: 940  
 Score: 1083.00 Matches: 190  
 Percent Similarity: 75.24% Conservative: 41  
 Best Local Similarity: 61.89% Mismatches: 76  
 Query Match: 13.31% Indels: 1  
 DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x AL550994 (1-940)

Qy 858 TrpAlaGluTyrPheGlnCysGluGlySerGluThrHisLeuAlaLeuCysProIleVal 877  
 Db 9 TGGGCTGAAGATCCAGTGTGAGGACATGAGTCCCTTTCACCTTCGTCYAGTACCA 68  
 Qy 878 GlnHisProGluAspThrCysIleHisSerArgValGlyValValCysSerArgTyr 897  
 Db 69 CCCCGCCAGAGAAGCACTGTAGCCACAGCAGGAGGTGTGGAGTAGTCTGCTCAAGTAC 128  
 Qy 898 ThrAspValArgLeuValAsnGlyLysSerClnCysAspGlyGlnValGluIleAsnVal 917  
 Db 129 ACAGAAATCCCTTGGTGAATGGCAAGACCCCGTGTGAGGACAGTGGAGCTCAAAACG 188  
 Qy 918 LeuGlyHisTrpGlySerLeuCysAspThrHisTrpAspProGluAspAlaArgValLeu 937  
 Db 189 CTGGTCCCTGGGATCCCTCTGTAACTCTCAGTGGACATAGAGATGCCCATGTCTT 248  
 Qy 938 CysArgGlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGlu 957  
 Db 249 TGCAGCAGCTTAAATGTGGAGTTCCTTCTACCCAGGAGCAGCAGCTTTTGAAAA 308  
 Qy 958 ArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAsp 977  
 Db 309 GGAATGGTCAGATGTGGAGCATATGTTTCACCTGCATGGAGTGGACAGCAGCATGGGA 368  
 Qy 978 AsnCysGlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerVal 997  
 Db 369 GATGTCTGTACTGTCTAGTGTCTCATATGCTTCCAGACAGTGGCCCTCTGTA 428  
 Qy 998 IleCysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspPro 1017  
 Db 429 ATCTGCTCAGGAAACCACTGCCAACACTGTCTCTGTCATTCATCTCTTTGGGCCCA 488  
 Qy 1018 TyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArg 1037  
 Db 489 ACAAGGCTTACCATTCCAGAAAGAGTGTGTGCTGCTCATAGAGAGTGTGTCACCTGC 548  
 Qy 1038 LeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrp 1057  
 Db 549 CTGTAATGGAGGAGTGGCTGTCTGGGAGTAGAGATCTATCATGAGGCTCTCTGG 608  
 Qy 1058 GlyThrIleCysAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeu 1077  
 Db 609 GGCACCATCTGTATGACAGCTGGACCTGAGTGTATGCCACGTGTTTCAGACAGCTG 668  
 Qy 1078 GlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyPro 1097  
 Db 669 GCCTGTGGAGAGCATTAAATGCCACTGGTCTCTCTATTGTTGGGGAAGCAACAGGCC 728  
 Qy 1098 IleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSer 1117  
 Db 729 ATCTGGCTGGATGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 788  
 Qy 1118 ArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGlu 1137  
 Db 789 CACGGCTGGGGCAGCAAAATTCAGGCACAAAGGAGTGGGGAGTATCTGCTCAGAA 848  
 Qy 1138 PheThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGlu 1157  
 Db 849 TTATGCTCTCTKAGACTGACCAAGTGAAGCCAGCAGAGAGGCTCTGCGAGGSGTCTGAA 908  
 Qy 1158 ValPheTyrAsnGlyThrTrp 1164  
 Db 909 GT-TTTTACATGAGCTTGG 928

RESULT 3  
 BG570972  
 LOCUS  
 DEFINITION  
 mRNA sequence.  
 BG570972  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 human.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

1. (bases 1 to 582)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps@email.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LICM1558 row: k column: 12  
 High quality sequence stop: 576.

FEATURES  
 source

1..582  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4714115"  
 /lab\_lib="NIH\_MGC\_79"  
 /lab\_host="DH10B (TI phase-resistant)"  
 /note="Organ: placenta; Vector: pDNR-LIB (Clontech);  
 Site\_1: SfII (ggccgctcgcc); Site\_2: SfII (ggccattatggcc);  
 5' and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGCGAGCGCGGCACATG-dT(30)BN-3',  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH-MGC  
 Library."

BASE COUNT 126 a 140 c 198 g 118 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.38e-98 Length: 582  
 Score: 1072.00 Matches: 192  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.48% Mismatches: 0  
 Query Match: 13.17% Indels: 0  
 DB: 12 Gaps: 0

US-09-759-130B-381 (1-1453) x BG570972 (1-582)

Qy 1109 GluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLys 1128  
 Db 2 GAGTCCCATCTGTGGCAGTGCCTTCCCGCGCTGGGGCAGCAGCAGTGCAGGCACAAG 61  
 Qy 1129 GluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGlu 1148  
 Db 62 GAGCAGCAGGGGTCATCTGCTCAGAAATTCACACCTTGAGGCTCTACAGTGAACACTGAA 121  
 Qy 1149 ThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValcyl 1168  
 Db 122 ACAGAGAGCTGTGCTGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCTGGC 181  
 Qy 1169 ArgArgAsnIleThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGlu 1188  
 Db 182 AGGAGGAACATCACACAGGCATAGCAGGCAATTGTGTGACAGCAGCTGGGCTGTGGGAG 241

BG570972 582 bp mRNA linear EST 10-APR-2001  
 602591546F1 NIH\_MGC\_79 Homo sapiens cDNA clone IMAGE:4714115 5',  
 mRNA sequence.

BG570972  
 BG570972.1 GI:13578625  
 EST.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1. (bases 1 to 582)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps@email.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LICM1558 row: k column: 12  
 High quality sequence stop: 576.

FEATURES  
 Location/Qualifiers  
 1..582  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4714115"  
 /lab\_lib="NIH\_MGC\_79"  
 /lab\_host="DH10B (TI phase-resistant)"  
 /note="Organ: placenta; Vector: pDNR-LIB (Clontech);  
 Site\_1: SfII (ggccgctcgcc); Site\_2: SfII (ggccattatggcc);  
 5' and 3' adaptors were used in cloning as follows: 5'  
 adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor  
 sequence: 5'-ATTCTAGAGCGAGCGCGGCACATG-dT(30)BN-3',  
 (where B = A, C, or G and N = A, C, G, or T). Average  
 insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies  
 contained inserts by PCR. This library was enriched for  
 full-length clones and was constructed by Clontech  
 Laboratories (Palo Alto, CA). Note: this is a NIH-MGC  
 Library."

BASE COUNT 126 a 140 c 198 g 118 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.38e-98 Length: 582  
 Score: 1072.00 Matches: 192  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.48% Mismatches: 0  
 Query Match: 13.17% Indels: 0  
 DB: 12 Gaps: 0

US-09-759-130B-381 (1-1453) x BG570972 (1-582)

Qy 1109 GluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLys 1128  
 Db 2 GAGTCCCATCTGTGGCAGTGCCTTCCCGCGCTGGGGCAGCAGCAGTGCAGGCACAAG 61  
 Qy 1129 GluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGlu 1148  
 Db 62 GAGCAGCAGGGGTCATCTGCTCAGAAATTCACACCTTGAGGCTCTACAGTGAACACTGAA 121  
 Qy 1149 ThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValcyl 1168  
 Db 122 ACAGAGAGCTGTGCTGGAGATTGGAAGTCTTCTATAACGGGACCTGGGGCAGCTGGC 181  
 Qy 1169 ArgArgAsnIleThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGlu 1188  
 Db 182 AGGAGGAACATCACACAGGCATAGCAGGCAATTGTGTGACAGCAGCTGGGCTGTGGGAG 241

QY 1189 AsnGlyValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValasp 1208  
|||||  
Db 242 AATGAGTGTGAGCTCGCCCTTATCTAAGACAGGCTCTGGTTTCATGTGGGTGGAT 301  
  
QY 1209 AspIleGlnCysProLysThrHisLeuSerIleTrpGlnCysLeuSerAlaProTrpGlu 1228  
|||||  
Db 302 GACATTCAGTGTCTAAACGATATCTCCATATGCGAGTGTCTGTGCCCCATGGGAG 361  
  
QY 1229 ArgArgIleSerSerProAlaGluThrTrpIleThrCysGluAspArgIleArgVal 1248  
|||||  
Db 362 CGAAGAATCTCCAGCCACAGACAGACCTGGATCATGTGAAGATCGAATAGAGTG 421  
  
QY 1249 ArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGly 1268  
|||||  
Db 422 CTGGAGGAGACACCGAGTGTCTGGGAGAGTGGAGATGTGCGACGAGCTCTGGGGC 481  
  
QY 1269 ThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnLeuGly 1288  
|||||  
Db 482 ACAGTGTGTGATGACTCTGGGACCTGGCCGAGGCGGAAGTGTGTGTCAGCAGCTGGGC 541  
  
QY 1289 CysGlySerAlaLeuAlaLeuArgAspAlaSerPhe 1301  
|||||  
Db 542 TGTGCTCTGCTCTGCTGCTGCTGAGGACGCTTCTGTAT 580  
  
RESULT 4  
BI185454  
LOCUS 799 bp mRNA linear EST 10-JUL-2001  
DEFINITION UNL-P-FN-cw-d-01-0-UNL-s1 UNL-P-FN Sus scrofa cDNA clone  
UNL-P-FN-cw-d-01-0-UNL 3', mRNA sequence.  
ACCESSION BI185454  
VERSION BI185454.1 GI:14659863  
KEYWORDS EST.  
SOURCE pig.  
ORGANISM Sus scrofa  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
Caetano, A.R., Johnson, R.K. and Pomp, D.  
Generation and sequence characterization of a normalized cDNA  
library from swine ovarian follicles  
Unpublished (2001)  
Contact: Pomp, D  
Department of Animal Science  
University of Nebraska, Lincoln  
Lincoln, NE 68583-0908, USA  
Tel: 402 472 6416  
Fax: 402 472 6362  
Email: dpomp@unl.edu  
Oligo-dT track not found, Not 1 site shown in beginning of sequence  
is likely internal to the message.  
Seq primer: M13 -29  
POLYA-No.

FEATURES  
source  
1..799  
/organism="Sus scrofa"  
/strain="University of Nebraska, Lincoln Swine Selection  
Lines"  
/db\_xref="taxon:9823"  
/clone="UNL-P-FN-cw-d-01-0-UNL"  
/dev\_stage="ADULT"  
/lab\_host="DHI0B (Life Technologies)"  
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not 1; Site 2: Eco RI; The UNL-P-FN  
library is a normalized library representing porcine  
ovarian follicles, ranging between 2.0 to 10.0 mm in  
diameter, collected during 7 days of the follicular phase  
of the pig estrous cycle. This library was derived from  
the library UNL-P-F2. The tag is a string of 5-6  
nucleotides present between the Not I site and the  
oligo-dT track. The library was constructed as described  
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806

, 1996.  
TAG\_SEQ=None found\*  
BASE COUNT 206 a 174 c 241 g 177 t 1 others  
ORIGIN  
  
Alignment Scores:  
Pred. No.: 2,31e-86 Length: 799  
Score: 960.00 Matches: 165  
Percent Similarity: 79.17% Conservative: 44  
Best Local Similarity: 62.50% Mismatches: 55  
Query Match: 11.80% Indels: 1  
DB: 13 Gaps: 0  
  
US-09-759-130B-381 (1-1453) x BI185454 (1-799)  
QY 999 CysThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyr 1018  
|||||  
Db 9 TCTCAGGAGACAGACAGACACTATCCCGCTGCAATTCATCATCTCGGACCCATCA 68  
  
QY 1019 LeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAspLysArgLeuArgLeu 1038  
|||||  
Db 69 AGCTCTATTATTCAGAAAGAAATGTTGCTGCTGCATAGGAGGTGGTCAACTTCGCTG 128  
  
QY 1039 ValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAspGlyPheTrpGly 1058  
|||||  
Db 129 GTCGATGGAGGTGGTGTGCTGGGAGAGTAGAGGTCTATCATGAGGGCTCTCGGGC 188  
  
QY 1059 ThrIleCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGly 1078  
|||||  
Db 189 ACCATCTGTGATGACAGCTGGGACCTGAATGATGCCATGTGTGTGCAACAGCAATAGC 248  
  
QY 1079 CysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProfile 1098  
|||||  
Db 249 TGTGATGGGCCATTATATGCCACTGTTCTGCTCATTTTGGGGAAGGAACAGGGCCATT 308  
  
QY 1099 TrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArg 1118  
|||||  
Db 309 TGGCTGGATGAGATAAATGTAATGGAAGAATCTCATATTGGCAATGCCACTCACAT 368  
  
QY 1119 GlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSerGluPhe 1138  
|||||  
Db 369 GGTGGGGGGCGCACAAATGTCAGGCATTAAGAGAGTAGCAGGAGTATCTGCTCGGAGTTC 428  
  
QY 1139 ThrAlaLeuArgLeuTyrSerGluThrGluThrGluSerCysAlaGlyArgLeuGluVal 1158  
|||||  
Db 429 ATGTCTCTGAGACTGATCAGTGAACACAGCAGAGAGACCTGTGCAGGGCGCTGGAAGTT 488  
  
QY 1159 PheTyrAsnGlyThrTrpGlySerValGlyArgArgAsnIleThrThrAlaIleAlaGly 1178  
|||||  
Db 489 TTTTACACGGAGCTGGGGCAGCGTTGGCAAGATAGCATGTCTCCAGCCACAGTGGGG 548  
  
QY 1179 IleValCysArgGlnLeuGlyCysGlyGluAsnGlyValValSerLeuAlaProLeuSer 1198  
|||||  
Db 549 GTGGTATCGAGCGCAGTAGGCTGTGCAGACAGAGGGACATCAGCCCTGCATCTTCAGAC 608  
  
QY 1199 LysThrGlySerGlyPheMetTrpValAspIleGlnCysProLysThrHisIleSer 1218  
|||||  
Db 609 AAGACAGTGTCCAGGCACATGTGGGTGGCAATGTTTCAGTGTCTCTAAAGGACCTGACACA 668  
  
QY 1219 IleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerProAlaGluGluThr 1238  
|||||  
Db 669 CTATGGCAGTGGCCATCTCCATGGGAAGAGAGACTGGCCAGCCCTCAGAGGAGACA 728  
  
QY 1239 TrpIleThrCysGluAspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArg 1258  
|||||  
Db 729 TGGATCATGTGCCCAACAAATAAGACTTCAAGAGGAAN-ACTAATTGTTCTTGACCT 787  
  
QY 1259 ValGluIleTrp 1262  
|||||  
Db 788 GTGGAGATCTGG 799  
  
RESULT 5  
AL544534

LOCUS AL544534 1108 bp mRNA linear EST 16-FEB-2001  
 DEFINITION AL544534 LTI\_NFL006\_PL2 Homo sapiens cDNA clone CS0D1021YH23 5  
 prime, mRNA sequence.  
 ACCESSION AL544534  
 VERSION AL544534.1 GI:12877014  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1108)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 FEATURES  
 Location/Qualifiers  
 1. .1108  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1021YH23"  
 /tissue\_type="placenta"  
 /notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"  
 BASE COUNT 307 a 189 c 319 g 289 t 4 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2 58e-84 Length: 1108  
 Score: 942.50 Matches: 169  
 Percent Similarity: 65.58% Conservative: 52  
 Best Local Similarity: 50.15% Mismatches: 111  
 Query Match: 11.58% Indels: 5  
 DB: 9 Gaps: 3  
 US-09-759-130B-381 (1-1453) x AL544534 (1-1108)  
 Qy 1 MetMetLeuProGlnAsnSerTrpHisIleAspPheGlyArgCysCysHis---Gln 19  
 Db 69 ATGGTGCTACTTGAAGACTCTGGATCTGCTGACTTCAGAGA-----CATTTGTC 119  
 Qy 20 AsnLeuPheSerAlaValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSer 39  
 Db 120 AACTTGAGTCCCTTCACATTACTGTGGTCTTCTCTAGTGGCTGTGTGTCACCACT 179  
 Qy 40 SerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGly 59  
 Db 180 TCTCTGGAGGAACAGACAGGAGCTGAGCTAGTGGATGGTGAACAACTGTAGCGNG 239  
 Qy 60 ThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspAspGlyTrpAsnThr 79  
 Db 240 AGAGTGAAGTGAAGTCCAGAGGAGTGGGAAACGGTGTGTAATAATGGCTGGAGCATG 299  
 Qy 80 ThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMetPheArg 99  
 Db 300 GAAGCGGTCTCTGTGATTTGTAACCACTGGGATGTCACACTGTATCAAGCCCTCGGA 359  
 Qy 100 PheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspValSerCysTrpGly 119  
 Db 360 TGGCTAATTCAGTGGAGGTTCTGGAGCAATTTGGATGGATCATGTTCTTGTCTGGGG 419  
 Qy 120 AsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHis---AsnCysTrp 138

Db 420 AATCAGTCAGCTCTTTGGGATGCAAAACATGATGGATGGGAAAGCATAGTAACGTACT 479  
 Qy 139 HisGlyGluAspValGlyValAsnCysTrpGlyGluAlaAsnLeuGlyLeuArgLeuVal 158  
 Db 480 CACCAACAAGATGCTGGGTGACCTGCTCAGATGGATCCCAATTTGGAAATGAGGCTGACG 539  
 Qy 159 AspGlyAsnAsnSerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThr 178  
 Db 540 CGTGGAGGATATGCTGTTGGGAAGATAGATCAAAATTCGAAGCAGCGTGGGGAACA 599  
 Qy 179 IleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCys 198  
 Db 600 GTGTGTGATGATAACTTCAACATAGATCATGCATCTGTCTTGTAGACAACTTGAATGT 659  
 Qy 199 ProSerSerPheIleSerSerGlyValValAsnSerProAlaValLeuArgProIleTrp 218  
 Db 660 GGAAGTCTGTGCTGCTTCTCTGTTTCATCTAAATTTGGAGAGGCTCTGGACCAATCTGG 719  
 Qy 219 LeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGly 238  
 Db 720 TTTGATGATCTTATATGCAACGGAATGAGTCAGCTCTCTGGAATGCAAAACATCAAGGA 779  
 Qy 239 TrpGlyAsnHisAspCysSerSerHisAsnGluAspValThrLeuThrCysTrpAspSer 258  
 Db 780 TGGGAAAGCATAACTGTGATCAGTCTCAGGATGCTGAGTGTGCTCAAGGAGGACA 839  
 Qy 259 AspLeuGluLeuArgLeuValGlyThrAsnArgCysMetGlyArgValGluLeuLys 278  
 Db 840 GATCTGAGCTGACGAGTGGTAGATGGAGTCACTGAATGTTCCAGGAAGATTAGAAGTGA 899  
 Qy 279 IleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAlaAspVal 298  
 Db 900 TTCAAGGAGAGTGGGAGACAATATGATCAGCGCTGGGACACTAGGATGCTGCTGTG 959  
 Qy 299 ValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGln 318  
 Db 960 CATGCAAGCAACTGGGATGTCCAACTCCGCTCAGCAGCAATGGTTCGAGTTAAGCCCAT 1019  
 Qy 319 SerGlySerAspValValTrpLeuAspClyValSerCysSerGlyAsnGlu 335  
 Db 1020 VAGGATTTGGACACATCTGGCTTGACAGCGCTTCTTCCAGGAGACATGAA 1070  
 RESULT 6  
 AL540905 1100 bp mRNA linear EST 16-FEB-2001  
 LOCUS AL540905 LTI\_FL002\_PL1 Homo sapiens cDNA clone CS0DE004YI19 5 prime  
 DEFINITION , mRNA sequence.  
 ACCESSION AL540905  
 VERSION AL540905.1 GI:12871463  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1100)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 FEATURES  
 Location/Qualifiers  
 1. .1100  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE004YI19"  
 /lab\_host="DH10B"  
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I



Pred. No.:	2,65e-83	Length:	718
Score:	929.00	Matches:	156
Percent Similarity:	82.13%	Conservative:	37
Best Local Similarity:	66.38%	Mismatches:	42
Query Match:	11.42%	Indels:	1
DB:	13	Gaps:	0

US-09-759-130B-381 (1-1453) x B1183796 (1-718)

Qy	1050	Glut1eYrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAsp	1069
Db	14	GAGTCTATCATGAGGGCTCTCGGGGACCACCTGTGTATGACACTGGGACCTGAATGAT	73
Qy	1070	AlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAla	1089
Db	74	GCCATGTGGTGTGCAACACAGCTGAGCTGTGGATGGGCCATTATGCCACTGGTCTGCT	133
Qy	1090	HisPheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGlu	1109
Db	134	CATTTTGGGAAGGAACAGGGCCCATTTGGCTGGATGAGATAAACTGTAATGGAAAGAA	193
Qy	1110	SerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysAtqHisLysGlu	1129
Db	194	TCTCATATTGGCAATGCCACTACATGTTGGGGGGCGGCAATTCAGGCATAAGGAG	253
Qy	1130	AspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuYrSerGluThrGluThr	1149
Db	254	GATCAGGAGTCATCTGCTGGAGTTCATGTCCTCAGACTGATCAGTGAACAGCAGA	313
Qy	1150	GluSerCysAlaGlyArgLeuGluValPheYrAsnGlyThrTrpGlySerValGlyArg	1169
Db	314	GAGACCTGTGCGGGCGCTGGNAGTTTTTTACAAACGAGCTTGGGGCAGCGTTGGCAAG	373
Qy	1170	ArgAsnIleThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsn	1189
Db	374	AATGACATGTCACAGCACAGTGGGGGTGTATGCGAGCAGCTGGGCTGTGCAGACAGA	433
Qy	1190	GlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp	1209
Db	434	GGGACATCAGCCCTGCATCTTCAGACAAGACAGTGTCCAGGCACATGTGGGTGGACAAT	493
Qy	1210	IleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArg	1229
Db	494	GTTCAGTGTCTTAAAGGACCTTGACACCCATATGGCAGTGCCTCATCTCCATGGAAGAAG	553
Qy	1230	ArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgValArg	1249
Db	554	AGACTGGCCACCCCTCAGAGGACATGGATCATGTGCCAACAAAATAAGACTTCAA	613
Qy	1250	GlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpGlyThr	1269
Db	614	GAAGNACACATAATTGTTCTNGACGTGTGGAGATCTGGTACGGNGTTCCTGGGGCACT	673
Qy	1270	ValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCys	1284
Db	674	GTGTGTACGACTCTCTGGGACTTGAA-AATGCNCAAGTGTGTGC	717

RESULT 8  
AL529438/c AL529438 554 bp mRNA linear EST 13-FEB-2001

DEFINITION AL529438 LTI\_NFL001\_NBC4 Homo sapiens cDNA clone CS0DD006Y005 3 prime, mRNA sequence.

ACCESSION AL529438

VERSION AL529438.1 GI:12792931

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 554)  
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
source  
1. .554  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="LTI\_NFL001\_NBC4"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"  
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 132 a 166 c 127 g 110 t 19 others

ORIGIN

Alignment Scores:  
Pred. No.: 4,57e-80 Length: 554  
Score: 895.00 Matches: 164  
Percent Similarity: 90.81% Conservative: 4  
Best Local Similarity: 88.65% Mismatches: 16  
Query Match: 11.00% Indels: 1  
DB: 9 Gaps: 0

US-09-759-130B-381 (1-1453) x AL529438 (1-554)

Qy	1189	AsnGlyValValSerLeuAlaProLeu-SerLysThrGlySerGlyPheMetTrpValAs	1208
Db	554	AATGGAGTTATCAGCTCGCCCTTTATTCTAAGACAGGCTCTGTTTCATGTGGGTGA	495
Qy	1208	PASpileGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpG	1228
Db	494	TRACATTCAATGTCTTAAACGCATGTCTCCATATGGCAGTGCCTSTCTGCCCATGGGA	435
Qy	1228	uArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArgVa	1248
Db	434	CGAAGASTCTCCAGCCCGCAGAGAGACCTGGCTMACATGTGAAGATAGATAGAGT	375
Qy	1248	lArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrpG	1268
Db	374	CGGTGGAGGAGACACCGAGTCTCTGGGAGAGTGGAGATCTGGCAGCAGGCTCTCTGGG	315
Qy	1268	yThrValCysAspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGluLeuG	1288
Db	314	CACAGTGTGTATVACTCTCTGGGACCTGGCGGAGCGAGNAGTGTGTNTCAGACGCTGG	255
Qy	1288	yCysGlySerAlaLeuAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThr	1308
Db	254	CTGTGGCTCTCTCTNGCTGCCCTGAGGAGCAGCTCTGTTTGGCCAGGGGACTGGACNCT	195
Qy	1308	eTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLy	1328
Db	194	CTGGTGGNTTACATGCGGTGCAAAAGGAANTNAGTCATTTCTATGGGACTGTGCACGCCA	135
Qy	1328	sProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyG	1348
Db	134	ACCTGGGAGACAGTACTGTGGACACAAAGGAGANGCTGGCGTGTAGTGTCTCTGGACA	75
Qy	1348	nSerLeuLysSerLeuAsnAlaSerSerGlyHisLeuAlaLeuLeuLeuSerSerIlePh	1368
Db	74	CTCGCTGNAATCAGTGAATGCTCTCTCAGGTCATTTAGGACTTAATTTATCAGTGTCTT	15
Qy	1368	eGlyLeuLeuLeu 1372	

```

||||| |||
Db 14 TGGCTCCKTCTC 2

RESULT 9
BI838644
LOCUS 930 bp mRNA linear EST 04-OCT-2001
DEFINITION 603086185F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5225205 5',
mRNA sequence.
ACCESSION BI838644
VERSION BI838644.1 GI:15950194
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 930)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11566 row: j column: 22
High quality sequence stop: 731.
Location/Qualifiers
1..930
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5225205"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT 236 a 207 c 274 g 213 t
ORIGIN

Alignment Scores:
Pred. No.: 1,07e-73 Length: 930
Score: 836.50 Matches: 163
Percent Similarity: 66.88% Conservative: 47
Best Local Similarity: 51.91% Mismatches: 98
Query Match: 10.28% Indels: 7
DB: 13 Gaps: 3

US-09-759-130b-381 (1-1453) x BI838644 (1-930)

QY 940 GlnLeuSerCysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSer 959
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 CAGCTTAATGTGGAGTGGCTTCTTACCCAGGAGGAGGACGCTTTGGAAAGGAAT 61
QY 960 ValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCys 979
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 GGTGAGATCTGGAGGCAATGTTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY 980 GlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSerValIleCys 999
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 CTGTAACTGCTCTAGGTCCTTCAATATGCTTCTGAGCAAGTGGCTCTCTAATCTGC 181
QY 1000 ThrGlySerLeuThrGlnProLeuPheProCysLeuAlaAsnValSerAspProTyrLeu 1019
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

/db\_xref="taxon:9606"  
 /clone="CS0DE004Y119"  
 /clone\_lib="LTI\_FL002\_PL1"  
 /lab\_host="DH10B"  
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand  
 cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-stranded cDNA was digested with Not I  
 and cloned into the Not I and Eco RV sites of the  
 pCMVSPORT 6 vector. Library was constructed by Life  
 Technologies. Contact: Feng Liang Life Technologies, a  
 division of Invitrogen 9800 Medical Center Drive Rockville  
 , Maryland 20850, USA Fax: (1) 301 610 8371 Email:  
 fliang@lifetech.com URL:  
 http://fulllength.invitrogen.com"

BASE COUNT 292 a 300 c 228 g 313 t 20 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.83e-66 Length: 1153  
 Score: 767.00 Matches: 146  
 Percent Similarity: 65.45% Conservative: 51  
 Best Local Similarity: 48.50% Mismatches: 88  
 Query Match: 9.42% Indels: 16  
 DB: 9 Gaps: 5

US-09-759-130B-381 (1-1453) x AL540904 (1-1153)

Qy 1164 TrpGlySerValGlyArgArgAsnIleThrAlaIleAlaGlyIleValCysArgGln 1183  
 Db 1153 TGGGGCAGCTGTGGCAAGAGCAATGCTGAAMCCACTGTGGGTGGTGGCAGGCG 1094  
 Qy 1184 LeuGlyCysGlyGluAsnGlyValSerLeuAlaProLeuSerLysThrGlySerGly 1203  
 Db 1093 CTGGCTGTGCAGCAAGGGAATATCACCCTGCTTGTAGACAAGCCATGTCATT 1034  
 Qy 1204 PheMetTrpValAspPileClnCysProLysThrHisIleSerIleTrpGlnCysLeu 1223  
 Db 1033 CCCATGTGGGTGGCAATGCTCAGTGTCCAAAGGACCTGCACACGTTGGCAGTGC 974  
 Qy 1224 SerAlaProTrpGluArgIleSerSerProAlaGluGluThrTrpIleThrCysGlu 1243  
 Db 973 TCATCTCCATGGGAGAGAGAGTGGCCAGCCCTCGGAGGAGACCTGGATCATGTGAC 914  
 Qy 1244 AspArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValClnIleTrpHis 1263  
 Db 913 AACAGATAAGACTTCAGGAAGGCCACTTCTCTTCTGGACGTTGGAGATCTGGCAT 854  
 Qy 1264 AlaGlySerTrpGlyThrValCysAspAspSerTrpAspLeuAlaGluAlaGluVal 1283  
 Db 853 GGAGGTTCTCTGGGACAGTGTGTGATGATTCTTGGGACTTGGACGATCTCAGGTGG 794  
 Qy 1284 CysGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGln 1303  
 Db 793 TGTCAACACTTGGCTGTGGTCCAGCTTAAAGCATTCACCAAGAGAGAGTTGGTCAG 734  
 Qy 1304 GlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrp 1323  
 Db 733 GGGACTGACCCGATATGCTCAATGAGTGAAGTCAAGGGAATGAGTCTCTCTTGTGG 674  
 Qy 1324 AspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyVal 1343  
 Db 673 GATTGCTCTGCAGACGCTGGGCCATAGTGTGAGTGTGGCAGCAAGAGACGCTCAGTG 614  
 Qy 1344 ArgCysSerGlyGlnSerLeuLys-----SerLeuAsnAlaSerSer 1357  
 Db 613 AATTGCACAGATATTTCAGTGCAGAAAAACCCACAAAAAGCCACACAGGTTCGCTATCC 554  
 Qy 1358 GlyHisLeuAlaLeuIleLeuSerSerIlePheGlyLeuLeuLeuValLeuPheIle 1377  
 Db 553 CGTCAGTCATCTTTATKGCAGTCGGGATCTTGGGTGTCTTGTGGCCATTTTCGTC 494  
 Qy 1378 -----LeuPheLeuThrTrpCysArgValGlnLysGlnLysHisLeuProLeuArgVal 1395

Db 493 GCATTATTCTTCTGACTAAAAAGCGAAGACAGACAGACGMS-----CTTGCA GTT 443  
 Qy 1396 SerThrArgArgArgGlySerLeuGluGluAsnLeuPheHisGluMetGluThrCysLeu 1415  
 Db 442 CCCCACAGAGAGAGAACTTAGTCCACCAATTCATACCGGAGATGAATCTTGCCTG 383  
 Qy 1416 LysArgGluAspProHisGlyThrArgThrSerAspThrProAsnHisGlyCysGlu 1435  
 Db 382 AAKGCAGATGATCTGGACCTAATGAATCTCTCAGAAAT-----TCCCATGAGTCAGCT 329  
 Qy 1436 AspAlaSerAspThrSerLeuLeuGlyVal-----LeuProAlaSerGluAlaThr 1452  
 Db 328 GATTTTCAGTCTGCTGGAACATAATTTCTGTGTCTAAATTTCTTCTTATTTCTGGAATGGA 269  
 Qy 1453 Lys 1453  
 Db 268 AAG 266  
 RESULT 11  
 BI836459  
 LOCUS 603082833F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5221905 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI836459  
 VERSION BI836459.1 GI:15948009  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 723)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11558 row: a column: 10  
 High quality sequence stop: 708.  
 FEATURES  
 Location/Qualifiers  
 1..723  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5221905"  
 /clone\_lib="NIH\_MGC\_120"  
 /lab\_host="DH10B"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."  
 BASE COUNT 173 a 160 c 237 g 153 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.24e-64 Length: 723  
 Score: 739.00 Matches: 144  
 Percent Similarity: 73.88% Conservative: 37  
 Best Local Similarity: 58.78% Mismatches: 54  
 Query Match: 9.08% Indels: 10  
 DB: 13 Gaps: 0  
 US-09-759-130B-381 (1-1453) x BI836459 (1-723)



Db 523 CTCAATGAAGTGAAGTGCACAGGGAATGAGTCTCTCTGTGGGATGCTCTGGCAGACGC 582

Qy 1330 TrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyVal 1343

Db 583 TGGGGCATAGTGTGGGACACAGGAAGACGCTCCAGTG 624

RESULT 13

LOCUS BQ327934

DEFINITION MR4-RT0048-070501-010-g04 RT0048 Homo sapiens cDNA, mRNA sequence.

ACCESSION BQ327934

VERSION BQ327934.1 GI:20945766

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 585)

Dias Neto, E., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Nagai, M.A., de Oliveira, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Goldman, G.H., Carvalhal, A.F., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, D.F., de Souza, S.J., and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR46t2-MR4-RT0048-070501-010-g04&t3=2001-05-07&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 24  
High quality sequence stop: 584.  
Location/Qualifiers  
1..585  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="RT0048"  
/dev\_stage="Adult"  
/note="Organ: kidney tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 144 a 122 c 186 g 133 t

ORIGIN

Alignment Scores:

Pred. No.: 2,52e-60 Length: 585

Score: 701.00 Matches: 122

Percent Similarity: 80.83% Conservative: 34

Best Local Similarity: 63.21% Mismatches: 37

Query Match: 8.61% Indels: 1

DB: 14 Gaps: 0

US-09-759-130B-381 (1-1453) x BQ327934 (1-585)

Qy 1128 LysGluAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThr 1147

Db 3 AGGGAGATCGGGAGTAC-TGCTCAATTCATGCTCTGAGACTGACCACTGAAGCC 61

Qy 1148 GluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerVal 1167

Db 62 AGCAGAGAGGCGCTGCGAGGCGTCTGGAAGTTTTTTTACAATGAGCTTGGGGCACTGTT 121

Qy 1168 GlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGly 1187

Db 122 GGCAGAGTAGCATGCTGAAACCACTGTGGGTGTGGTGTGCAGGCAGCTGGCTGTGCA 181

Qy 1188 GluAsnGlyValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpVal 1207

Db 182 GACAAAGGGAATCAACCTTCATCTTTAGACAAGGCCATGTCATCCATCCATGTTGGGTG 241

Qy 1208 AspAspIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrp 1227

Db 242 GACAATGTTCAAGTCTCAAAAGGACCTGACACGCTGTGGCAGTGCCCATCATCTCCATGG 301

Qy 1228 GluArgArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArgIleArg 1247

Db 302 CAGAAGAGACTGGCCAGCCCTCGGAGGAGACCTGGATCATGTGACAACAAGATAAGA 361

Qy 1248 ValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAlaGlySerTrp 1267

Db 362 CTTCAGGAAGACCCACTTCTCTGTCTGGACGTGTGGAGATCTGCAATGGAGGTTCTCTGG 421

Qy 1268 GlyThrValCysAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeu 1287

Db 422 GGGACAGTGTGTGATGACTCTTGGGACTTGGACCATGCTCAGGTGGTGTGCAACAACCT 481

Qy 1288 GlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGlyThrGlyThr 1307

Db 482 GCGTGTGGTCCAGCTTTGAAAGCATTCAAAGAAGCAGAGTTTGGTCAGGGGACTGGACCG 541

Qy 1308 IleTrpLeuAspMetArgCysLysGlyAsnGluSer 1320

Db 542 ATATGGCTCAATGAAGTGAAGTGCACAGGGAATGAGTCT 580

RESULT 14

LOCUS BE502724

DEFINITION h278d12.x1 NCI-CGAP Lu24 Homo sapiens cDNA clone IMAGE:3214103 3', similar to TR:Q07898 Q07898 M130 ANTIGEN PRECURSOR. ;, mRNA sequence.

ACCESSION BE502724

VERSION BE502724.1 GI:9705132

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 653)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapsb-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov  
Seq primer: -40UP from Glibco  
High quality sequence stop: 483.  
Location/Qualifiers  
1..653  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3214103"  
/clone\_lib="NCI-CGAP\_Lu24"  
/tissue\_type="carcinoid"



Qy 1245 ArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAla 1264  
:::|||||::: ||| ||| ||||||||| ||||||||| ||||||||| |||||||||  
Db 348 AAGATAAGACTTCAGGAAGGACCCACTTCCTGTTCTGGACGTGTGGACATCTGGCATGGA 289  
Qy 1265 GlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaValValCys 1284  
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||  
Db 288 GGTTCCTGGGGACAGTGTGTGATGACTCTTGGGACTTGGGACGATGCTCAGGTGGTGTGT 229  
Qy 1285 GlnGlnLeuGlyCysGlySerAlaLeuAlaAlaLeuArgAspAlaSerPheGlyGlnGly 1304  
||||| ||||||||| |||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 228 CAACAACCTGGCTGTGGTCCAGCTTTGAAGCATTCAAAGAGCAGAGTTTGGTCAGGGG 169  
Qy 1305 ThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeuTrpAsp 1324  
||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||  
Db 168 ACTGACCGGATATGGCTCAATGAAGTGAAGTCAAAAGGAATGAGTCTTCCTTGTGGGAT 109  
Qy 1325 CysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArg 1344  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 108 TGTCTGCCAGACGCTGGGGCCATAGTAGTGTGGGCACAAAGGAAGACGCTGCAGTGAAT 49  
Qy 1345 CysSerGlyGln 1348  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 48 CGCAGGTAG 37

Search completed: May 12, 2003, 17:31:28  
Job time : 2883 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2003, 10:35:25 ; Search time 79 Seconds  
(without alignments)  
5640.523 Million cell updates/sec

Title: US-09-759-130B-381

Perfect score: 8138

Sequence: 1 MMLPQNSWHIDFGRCCHQON.....CEDASDTSLLGLVPASEATK 1453

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09759130/runat\_06052003\_123555\_23377/app\_query.fasta\_1.1607  
-DB=Issued\_Patents\_NA -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09759130.cgn.1.1.41@runat\_06052003\_123555\_23377 -NCPD=6 -ICPD=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6 -FCAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2435	29.9	5802	4	US-09-341-587-4
2	1142.5	14.0	4360	1	US-08-470-350B-1
3	785.5	9.7	29598	4	US-09-341-587-6
4	715.5	8.8	2001	4	US-09-341-587-2
5	631.5	8.5	2178	3	US-09-034-916-1
6	634.5	7.8	2262	3	US-09-276-400-3
7	634.5	7.8	2262	4	US-09-448-076-3
8	634.5	7.8	2262	3	US-09-702-572-3
9	634.5	7.8	2920	3	US-09-276-400-1
10	634.5	7.8	2920	4	US-09-448-076-1
11	634.5	7.8	2920	4	US-09-702-572-1
12	618.5	7.6	28720	4	US-09-341-587-7

13	314	3.9	2285	1	US-08-477-674-9	Sequence 9, Appli
14	314	3.9	2285	1	US-08-473-791-9	Sequence 9, Appli
15	314	3.9	2285	2	US-08-316-714-9	Sequence 9, Appli
16	314	3.9	2285	3	US-08-473-673-9	Sequence 9, Appli
17	290.5	3.6	1868	1	US-08-392-367B-1	Sequence 1, Appli
18	290.5	3.6	1868	3	US-08-893-467A-1	Sequence 1, Appli
19	290	3.6	2028	1	US-08-453-117-1	Sequence 1, Appli
20	290	3.6	2028	2	US-08-948-222-1	Sequence 1, Appli
21	290	3.6	2028	2	US-08-973-145-1	Sequence 1, Appli
22	290	3.6	2028	2	PCT-US96-08081-1	Sequence 1, Appli
23	290	3.6	2037	1	US-08-134-365-1	Sequence 1, Appli
24	288.5	3.5	1560	2	US-08-794-795-5	Sequence 5, Appli
25	288.5	3.5	1560	4	US-08-249-200-5	Sequence 5, Appli
26	284.5	3.5	1703	2	US-08-794-795-1	Sequence 1, Appli
27	284.5	3.5	1703	4	US-09-249-200-1	Sequence 1, Appli
28	282	3.5	410	4	US-09-221-298-83	Sequence 83, Appli
29	268	3.3	1588	6	5510466-3	Patent No. 5510466
30	248.5	3.1	9723	1	US-08-083-590A-21	Sequence 21, Appli
31	248.5	3.1	9723	3	US-08-532-384-21	Sequence 21, Appli
32	244	3.0	3955	4	US-09-214-278-4	Sequence 4, Appli
33	243.5	3.0	4315	3	US-08-882-046-3	Sequence 3, Appli
34	243.5	3.0	8257	4	US-09-484-970B-65	Sequence 65, Appli
35	239.5	2.9	4483	3	US-08-611-729A-7	Sequence 7, Appli
36	239	2.9	4464	2	US-08-400-159-7	Sequence 7, Appli
37	208.5	2.6	5804	4	US-09-369-364A-12	Sequence 12, Appli
38	207	2.5	7286	5	PCT-US95-11684-1	Sequence 1, Appli
39	200	2.5	6677	4	US-08-939-366-27	Sequence 27, Appli
40	200	2.5	6677	4	US-09-467-997-6	Sequence 6, Appli
41	196.5	2.4	4394	2	US-08-750-152A-1	Sequence 1, Appli
42	133.5	2.4	16382	4	US-08-718-388-8	Sequence 8, Appli
43	186	2.3	4208	4	US-09-214-278-6	Sequence 6, Appli
44	185	2.3	5499	3	US-08-479-722B-1	Sequence 1, Appli
45	185	2.3	5502	5	PCT-US95-02251-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1

US-09-341-587-4  
; Sequence 4, Application US/09341587  
; Patent No. 6346606  
; GENERAL INFORMATION:  
; APPLICANT: Mollenhauer, Jan  
; TITLE OF INVENTION: Protein Containing an SRCR Domain  
; FILE REFERENCE: 4121-108  
; CURRENT APPLICATION NUMBER: US/09/341.587  
; CURRENT FILING DATE: 1999-08-31  
; EARLIER APPLICATION NUMBER: PCT/DE98/00096  
; EARLIER FILING DATE: 1998-01-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 5802  
; TYPE: ENA  
; ORGANISM: Homo sapiens  
US-09-341-587-4

Alignment Scores:

Pred. No.: 1.93e-215 Length: 5802  
Score: 2435.00 Matches: 533  
Percent Similarity: 46.58% Conservative: 155  
Best Local Similarity: 39.09% Mismatches: 415  
Query Match: 29.92% Indels: 374  
DB: 4 Gaps: 28

US-09-759-130B-381 (1-1453) x US-09-341-587-4 (1-5802)

Qy 43 GlyThrAsp-----LeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThr 60

Db 389 GGATCTGATTCTGGTTGGCCCTGAGCTGGTGAATGGAGATGGCAGGTGTCTAGGCCCA 448

Qy 61 ValGluValLysPheGlnGlnTrpGlyThrValCysAspAspGlyTrpAsnThr 80

Db 449 GTGGAGATCCTATACAGAGCTCTCTGGGACACCGTGTGTGTGATGACAGCTGGGACACCAAT 508  
Qy 81 AlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAlaMet----- 97  
Db 509 GATGCCAAGCTGGTCTGTAGGACAGTGGTGTGGCTGGCCATGTCACGCTCCAGGAAT 568  
Qy 98 PheArgPheGlyGlnAlaValThrArgHisGlyLysIleThrLeuAspValSerCys 117  
Db 569 GCCTGGTTTGGCCAGGC-----TCAGGACCAATGGCTGTGATGATGGGCTGC 619  
Qy 118 TyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpCysHisAsnCys 137  
Db 620 TCAGGACACCAATCTTACCTGTGGAGCTGCCCCCAACAATGGCTGGCTCTCCCACTACTGT 679  
Qy 138 TyrHisGlyGluAspValGlyValAsnCysTyr----- 148  
Db 680 GGCCATGGTGAAGATGCTGGTGTATCTGCTCAGCTGCCAGCCTCAGTCAACACTCAGG 739  
Qy 149 -----GlyGluAlaAsn 152  
Db 740 CCAAGAAAGTTGGCTGTCAAGGATATCACCACCTGTATCCACAGAGGATCTGAATCCAGT 799  
Qy 153 LeuGlyLeuArgLeuValAspGlyAsnAsnSerCysSerGlyArgValGluValLysPhe 172  
Db 800 TTGGCCCTCAGGCTGGTGAATGAGCGCACAGGTGTCGAGGCGGAGTGGAGGTCTCTATAC 859  
Qy 173 GlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValVal 192  
Db 860 CGAGGCTCTGGGACACCGTGTGTGATGACTACTGGGACACCAATGATGCCATGTGGTTC 919  
Qy 193 CysArgGlnLeuGlyCysProSerPheIleSerSerGlyValValAsnSerProAla 212  
Db 920 TGCAGGACGCTGGCTGTGGCTGGCCATGTCAGGCCAGGAATGCCAGTTTGGCCAG 979  
Qy 213 ValLeuArgProIleTrpLeuAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrp 232  
Db 980 GGCTCAGGACCCCATTTCTGGATGATGTGGCTGTCTCAGGACAGGCTCTACCTGTGG 1039  
Qy 233 AsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHisAsnGluAspValThrLeu 252  
Db 1040 AGCTGCCCCACAATGGCTGGCTCACCCACAACCTGGCCATAGTAGAAGACGCTGGTGC 1099  
Qy 253 ThrCysTyr----- 255  
Db 1100 ATCTGCTAGCTCCCGAGTCCCGGCGACACCCAGCCAGATCTTGGCCGACCTCACAT 1159  
Qy 256 -----AspSerSerAspLeuGluLeuArgLeuValGlyThrAsnArg 270  
Db 1160 GCATCAACAGCAGGACCTGAATCCAGTTTGGCCCTCAGGCTGGTGAATGAGGTGACAGG 1219  
Qy 271 CysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrValCysHisLys 290  
Db 1220 TGTGAGGCGGAGTGGAGTCTTATACCGAGGCTCTCTGGGACCCGCTGTGTGATGATAGC 1279  
Qy 291 TrpAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGlyThrAlaLeuHis 310  
Db 1280 TGGGACACCAAGTACGCAATGTGGTCTGCCGCACTGGCTGTGGCTGGGCGACGTCAC 1339  
Qy 311 PheAlaGlyLeuProHisLeuGlnSerGlySerAspValValTrpLeuAspGlyValSer 330  
Db 1340 GCCCCAGGAATGCCCGTTTGGCCAGGTTTACGAGCCCATTTCTGGATGACGTGGCG 1399  
Qy 331 CysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSerGlyThrValAsnAsp 350  
Db 1400 TGCTCAGGCTATGAGTCTACCTGTGAGCTGGCCCAACCAATGGCTGGCTCTCCCATAC 1459  
Qy 351 CysLeuHisGlnAsnAspValSerValIleCysSerAspGlyAlaAspLeuGluLeuArg 370  
Db 1460 TGTACACACAGTGAACACGCTGTGTCTATCTGCTCAGCTGCC----- 1501  
Qy 371 LeuAlaAspGlySerAsnAsnCysSerGlyArgValGluValArgIleHisGluGlnTrp 390  
Db 1502 -----CACTCCTGG 1510

Qy 391 TrpThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeu 410  
Db 1511 TCGAGC----- 1516  
Qy 411 GlyCysProPheSerValPheGlySerArgAlaLysProSerAsnGluAlaArgAsp 430  
Db 1517 -----CCAGTCCAGACACATGGCG 1537  
Qy 431 IleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysThrTyr 450  
Db 1538 ACATCACCCTTCCCTGCATCGACAGTAGGATCTGAATCCAGT----- 1579  
Qy 451 AspGlyLysAlaLysArgThrCysPheArgArgSerAspAlaGlyValIleCysSerAsp 470  
Db 1579 ----- 1579  
Qy 471 LysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTyrGlyArgLeuGlu 490  
Db 1580 -----TTGGCCCTGAGGCTGTAATGGAGGTGACAGGTGTCAGGCGCAGTGGAG 1630  
Qy 491 ValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSerThrArgAsnAla 510  
Db 1631 GTCTATACCAAGCTCTCTGGGACCGTGTGGATGACAGCTGGGACACCAATGATGCC 1690  
Qy 511 AlaValValCysLysGlnLeuGlyCysGlyLysProMetHisValPheGlyMetThrTyr 530  
Db 1691 AATGCTGTGAGGCAACCGGCTGTGGCTGGCCATGTGACGCCCCAGGAATGCCCG 1750  
Qy 531 PheLysGluAlaSerGlyProIleTrpLeuAspAspValSerCysIleGlyAsnGluSer 550  
Db 1751 TTTGGTCAAGGCTCAGGACCCCATTTCTGGATGATGTGGCTGCTCAGGACACGAGTCT 1810  
Qy 551 AsnIleTrpAspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgGluAsp 570  
Db 1811 TACCCGTGGAGCTGCCCCACAAATGGCTGCTCTCCCAACTGTGCCATAGTGAAGAC 1870  
Qy 571 ValIleValThrCysSerGlyAspAla-----ThrTrp--- 581  
Db 1871 GCTGGTGTCTCTGCTCAGCTTCCAGTCCCGGCCAACACCTAGTCCAGACACTTGGCCA 1930  
Qy 582 -----GlyLeuArgLeuValGly 588  
Db 1931 ACTCATCATCATCAACAGGAGTCTGAATCAGCTTGGCCCTGAGGCTGGTGAATGGA 1990  
Qy 589 SerAsnArgCysSerGlyArgLeuGluValThrPheGlnGlyArgTrpGlyThrValCys 608  
Db 1991 GGTGACAGGTGTGAGGCGGAGTGGAGTCTTATACCGAGGCTCTCTGGGACACCGTGT 2050  
Qy 609 AspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnLeuAspCysProSer 628  
Db 2051 GATGACTACTGGGACACCAATGATGCAATGTGGTTTGCAGGCGAGTGGCTGTGGCTGG 2110  
Qy 629 SerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyLysIleTrpLeuAsp 648  
Db 2111 GCATGTACCCCGGAGAAATGCCGGTTGGCCAGGTTTCAGGACCATTTCTCTGGAT 2170  
Qy 649 AspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArgAsnSerGlyTrpGly 668  
Db 2171 GATGTGGCTGTGTCAGGACATGATCTATCTGTGGAGTGGCCCCACCAATGGCTGCTC 2230  
Qy 669 AsnAsnAspCysSerHisSerGluAspValGlyValIleCysSer----- 683  
Db 2231 TCCCAACAATGTGGCCATCATGAAGACGCTGGTGTCTATCTGCTCAGCTTCCCCAGTCCAC 2290  
Qy 684 -----AspAlaSer 686  
Db 2291 CCGACACCCAGCCGACACTTGGCCCACTCATCATCATCAACAGCAGGATCTGAATCC 2350  
Qy 687 AspMetGluLeuArgLeuValGlyLysSerArgCysAlaGlyLysValGluValAsn 706  
Db 2351 AGTTTGGCCCTGAGGCTGGTGAATGGAGGTGACAGGTGTGAGGCGGCGGAGTGGAGTCTTA 2410





Db	1192	TATCTGTGGAGCTGCTCCACCAGAGCTGGCTCTCTATAACTGTGCCACCATGAGGAT	1251
Qy	678	ValGlyValIleCysSerAspAlaSer	688
Db	1252	GCTGGAGTCATCTGTTTCAGATGCCCAACCCAGACCAACCTGGCCAGATATGTGGCCT	1311
Qy	689	-----Glu-LeuArgLeuValGlyGlySerSerArgCysAlaGlyLysValG1	704
Db	1312	ACTTACCACCTCCAGAAACTACAACAGATGGTGGACT-ACAAAATATTCTCTCTGTCTCC	1370
Qy	704	uValasnValGlnGlyAlaValGlyLeuCysAlaAsnGlyTrpGlyMetAsnIleAl	724
Db	1371	TACAACACAAATTCCTCCACCATTA	1418
Qy	724	aGluValValCysArgGlnLeuGluCysGlySerAlaIleArgValSer	740
Db	1419	GGAAATACACC-----TGTTGGAGTTTACTGACCTACCTTATGGCAGTT	1463
Qy	741	-ArgGluProHisPheThrGluArgThrLeuHisIleLeuMetSerAsnSerGlyCysTh	760
Db	1464	TTCCAGCCCATACCTACCTCGGAAGC-----TATCCTAACAAATGCCAGATGT--	1509
Qy	760	rGlyGlyGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHi	780
Db	1510	-----TTGTGG-----	1515
Qy	780	sLeuAsnMetGluAlaSerLeuIleCysSerAlaHisArgGlnProArgLeuValGlyAl	800
Db	1516	-----AAATTTTCGTCTCCAGCATGAACCGTGTGACAGTGGTCTTCAC	1559
Qy	800	aaspMetProCysSerGlyArgValGluValLysHisAlaAspThrTrpArgSerValCy	820
Db	1560	AGATGTGCAGCTTGAAGGAGGT-----	1583
Qy	820	saspSerAspPheSerLeu-----HisAlaAlaAsnValLeuCysAr	834
Db	1584	CAACTATGACTACATCCTCGTTTGTATGGCCCTGAAACAAATTTCTCTCTCAPTGTCTG	1643
Qy	834	qGluLeuAsnCysGlyAspAlaIleSerLeuSerValGlyAspHisPheGlyLysGly--	853
Db	1644	G-----GTTTGTGATGGTTCATGGATCTTT	1670
Qy	854	-----AsnGlyLeuThrTrpAlaG1	860
Db	1671	CACCTCAACCCAGAACTTCATGCTGTAGTCTTTATCAGCATGGCAGTGTGCAGGAG	1730
Qy	860	uLysPheGlnCysGlu-----GlySerGluThrHisIleuAlaLe	873
Db	1731	AGGTTTCCAAAGCTGACTACTACTCCACTCTATCAGCACCAGCAACCTCTCCAACGAC	1790
Qy	873	uCysProIleVal-----GlnHisProGluAspThrCys-----	885
Db	1791	GTTCCTCCGATCGTTACTGATGTGGTGACAACTCTCTCCGGAATACACCTGTGGAGGTTT	1850
Qy	885	eHisSerArgGluValGlyValValCysSerArgTyrThrAspValArgLeuValasng1	905
Db	1851	ACTGACCTTACCCTATGGCGAGTTTTCAGGCCATAC-----	1887
Qy	905	yLysSerGlnCysAspGlyGlnValGluIleAsnValLeuGlyHisTrpGlySerLeuCy	925
Db	1888	-----TACCTTGGAGCTAT--	1902
Qy	925	saspThrHisTrpAspProGluAspAlaArgValLeuCysArg-----	939
Db	1903	-----CCTAACAAATGCCAGATGTTTGTGAAAAATTTTCGTCCCCAGCAT	1946
Qy	940	-----GlnLeuSerCysGlyThr-----	945
Db	1947	GAACCGTGTGACAGTGGTCTTCACAGATGTGCAGCTTGAACGAGGTTGCCAACTACTGACTA	2006
Qy	946	-AlaLeuSerThrThrGlyGlyLysTyrIleGlyGluArgSerValArgValTrpGlyHi	965

Db	2007	CATCCTGGGTTTGTAGGGCTGGAATACAAATCTCTCTCTCATTTGCTCGGGTT-----	2050
Qy	965	sArgpheHisCysLeuGlyAsnGluSerLeuLeuAspAsnCysGlnMetThrValLeuGlu	985
Db	2059	-----TGTGATGGGTCCAATGGATCTTTTCACTCAACCCAG-----	2094
Qy	985	yAlaProProCysIleHisGlyAsnThrValSerVal-----IleCysThrGlySerLe	1003
Db	2095	-----AACTTCATGCTGTGTAGTCTTTATPCACGGATGGCAGTGT	2132
Qy	1003	uThrGlnProLeuPhe-----ProCysLeuAlaAsnValSerAs	1016
Db	2133	CACGAGGAGGGTTCAGAGTCAGTACTTCCACTCTATCAGGACCAGCAACTCC	2192
Qy	1016	pProTyLeuSerAlaValProGluGly-----SerAlaLeuIleCysLeuGluAspLy	1034
Db	2193	TCCAACGAGCTCCGATCATTAATCTCGAAATGATCTTCATTTGGTG-----	2238
Qy	1034	sArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHisAs	1054
Db	2239	---CTGAGGCTGGTAATGGAACAACCGGTGTGAGGGCCGAGTGGAGATCTTGTACAG	2294
Qy	1054	pGlyPheTrpGlyThrIleCysAspGlyTrpAspLeuSerAspAlaHisValValCy	1074
Db	2295	AGGCTCTGGGTACCGTCTCCGACGACAGCTGGGACATCAATGATGCATATGGTCTG	2354
Qy	1074	sGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluG	1094
Db	2355	CAGACAGCTGGTGTGGCTCTCTCTGCTCTCCAGGAATGCTGTGTTTGGTTCAGGG	2414
Qy	1094	ySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpG	1114
Db	2415	TTCAGGCTCATTTGCTCGTGATGATGCTGCTCTCTGGGTATGATGATCCCACTGTG	2474
Qy	1114	nCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyVal	1134
Db	2475	TGTGCTGACCCCTGGCTGGCTTTCATATGTCGTCTGATGTTGAGGATGCAGGAGTCA	2534
Qy	1134	eCysSer-----GluPheThrAlaLeuArgLeuTySerGluThrGluThrGluSerCy	1152
Db	2535	TGTCTACTCCCTGATCCGACTCC-CTCTCTGGTCCAGTTTGGACAAGTCTCTCTTTG	2593
Qy	1152	sAlaGlyArgLeuGluValPheTyAsnGlyThrTrpGlySerValGlyArgArgAsnI	1172
Db	2594	T-----AAACTACTATTGTGGAGGTTCCTGACTGGAC-TCTCTGGCAATTTTCTAGC	2646
Qy	1172	eThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGlyVal	1192
Db	2647	CCATCACTACCCCTGGGAGCTATCTCTAATATGCCAGATGTTTGTGGA-----	2692
Qy	1192	lSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAsp-----	1209
Db	2693	-----ACATTGAAGTCCCAACAACTACCGCGTGACGTGCTTCCAGAGATGGCAG	2745
Qy	1210	-----IleGlnCysProLysThr-----HI	1216
Db	2746	CTGAAGGGGCTGCAACTATGACTATATAGAGATTTTTGATGGCCGCCACCACTTCA	2805
Qy	1216	sIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArgIleSerSerPro	1234
Db	2806	CTCTCATTTGCCGGGTTTGTGATGGGCCATGGGCTCTTTCACTTCAACATCCA	2860
RESULT 3			
US-09-341-587-6			
; Sequence 6, Application US/09341587			
; Patent No. 6346606			
; GENERAL INFORMATION:			
; APPLICANT: Mollenhauer, Jan			
; TITLE OF INVENTION: Protein Containing an SRCR Domain			
; FILE REFERENCE: 4121-108			
; CURRENT APPLICATION NUMBER: US/09/341,587			
; CURRENT FILING DATE: 1999-08-31			
; EARLIER APPLICATION NUMBER: PCT/DE98/00096			

; EARLIER FILING DATE: 1998-01-09  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 29598  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-341-587-6

## Alignment Scores:

Pred. No.: 3,896-61 Length: 29598  
 Score: 785.50 Matches: 501  
 Percent Similarity: 26.4% Conservative: 185  
 Best Local Similarity: 19.32% Mismatches: 520  
 Query Match: 9.65% Indels: 1390  
 DB: 4 Gaps: 103

US-09-759-130b-381 (1-1453) x US-09-341-587-6 (1-29598)

QY 35 CysPheLeuIleSerSerPheAsnGlyThrAspLeuGluLeuValAsnGlyAsp 54  
 DB 3837 TGCCCTCCCTAGGATCTGATCTGGT-----TTGGCCCTCAGGCTGGTGAATGAGAT 3890  
 QY 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74  
 DB 3891 GGCAGGTGTACAGGCGGAGTGGAGATCCTATACCGAGGCTCCTGGGACCCGCTGTGTGAT 3950  
 QY 75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94  
 DB 3951 GACAGCTGGGACCAACATGATCCCAACGTGTCTGTAGGCAGCTGGTGTGGCTGGGCC 4010  
 QY 95 PheAlaMet-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrp 111  
 DB 4011 ATGTGAGTCCAGGAATAACCTGGTGTGGCCAGGC-----TCAGGACCCATGTGCC 4061  
 QY 112 LeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGlu 131  
 DB 4062 CTGGATGATGTCGCTGCTCAGGACACGATCTTACCTGTGGAGCTGCCCCCAACATGCC 4121  
 QY 132 TrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCysTyrGlyGluAla 151  
 DB 4122 TGCTCTCCCAATACTGTGGCCATGCTGAAGATGCTGTGTATCTGCTCAGGTAGG--- 4178  
 QY 152 AsnLeuGlyLeuArgLeuValAspGlyAsnAsnSerCys----- 164  
 DB 4179 CATCCAGATCTCTGGAGGTTGGTGTGGTGGCTCATGCTTTAATCCCCACACTTTGGG 4238  
 QY 165 -----SerGlyArgValGluValLysPheGlnGluArgTrp 176  
 DB 4239 AGGATGAGTAGGCAGATTGCTTGAGCTCAGGCGTTCAAGACCACTCTAGGCCAAGATGG 4298  
 QY 177 GlyThr----- 178  
 DB 4299 CAACCCCATCTCTATTAAAAAGAAAATCACTTTGGGCCCATTTCTGGGACAGACTGTA 4358  
 QY 178 ----- 178  
 DB 4359 TTCAGAGATATAGTAACAGGTTTCTCACTCCAGGCGCCCGCCCAAGTCTTTCTCGTGT 4418  
 QY 179 -----IleCysAspAspGlyTrp 184  
 DB 4419 ATCTGTGGCTCTGAGTTTAGTTGGGCGAGTTGGCCCTGTGGGTACAAATGCCACGGTCAG 4478  
 QY 185 AsnLeuAsnThrAlaAlaValValCysArgGlnLeuGly----- 197  
 DB 4479 CACAATCCCAAGCTGCAGAGCGTGGCCGCAACATTTGTTGGAGGTGACTGAGGCACCA 4538  
 QY 197 ----- 197  
 DB 4539 GTGTTACAGACAGGTCAGCATGCTGACATCAGAAATGATGACCTTTTATCTGCTC 4598  
 QY 197 ----- 197

DB 4599 CAGGAACATCCGCATAAATCATACCCCATTTTCCCGAGGCTCTCTCTCAGGACAGCA 4658  
 QY 198 -----CysProSerSerPhe 202  
 DB 4659 CGTTTTTGGGATCTGGTGGTGAAGCCCTTCCCTCCCTTGAACACTGTATTACCTTGAGT 4718  
 QY 203 IleSerSer-----GlyValValAsnSerProAlaValLeu--- 214  
 DB 4719 GTTGTGAGCCGCTGACCCAGAGGAAGGATGCAATTGTAAATAGTTCACTTATGATTGCT 4778  
 QY 215 -----ArgProIle-----TrpLeuAspAspIleLeuCysGlnGlyAsn 227  
 DB 4779 GTGAAGAGAGGAGGAGGAGTGGGTAAAGGTGGGAGGAT-----GGCATT 4823  
 QY 228 GluLeuAlaLeu-TipAsn-----CysArgHisArgGlyTrpGlyAsnHisAspCys 244  
 DB 4824 GAGTCGCAATGGTGGGCTCACAGTACAGTACAGTACAGGAGGAGGAGGAGCTCCTAGA 4883  
 QY 244 sSerHisAsnGluAsp----- 249  
 DB 4884 GTATCACCAGGAATCCAGCTCAAACTCAGACTCATCGCAGCTCAAGCGTGGAGAGTTAG 4943  
 QY 250 -----ValThrLeuThrCysTyrAspSerSerAspLeuGlu 261  
 DB 4944 GGTCTTCATCTACTCATTCGTTAGGGTGGCTGTTTGTCTCTCCCTGGGTGTTGATACCGA 5003  
 QY 261 uLeuArgLeuValGly----- 266  
 DB 5004 GGTGTTGTGGCGGTGCCAGGACCGAAGTCACTTATCCCTGGGAGAGTGAGCCAC 5063  
 QY 267 GlyThrAsnArgCysMet----- 272  
 DB 5064 AGGCACAGAGAAGCAATGCCAGGCGTGTGGGTCCCTGAGGAGCCGAGGCTTCCACGG 5123  
 QY 273 -----GlyArgValGlu-LeuLysIleGlnGly----- 281  
 DB 5124 TGGCAGCTGGCAGGCGCCACACTTGGTGGCGTGTGATGGGCATCAGCTCAGGGGTAGATA 5183  
 QY 281 ----- 281  
 DB 5184 CCCCAGTCACTTTCAGCTTAACTCTACTTGGAGTCACTGAGTGTGGTGTCTAATGTT 5243  
 QY 281 ----- 281  
 DB 5244 GCTATTTTTTCTCAGAGTGGCCAGCTCAGTCAACACTCAGGCCAGGTGAGTCCCCAG 5303  
 QY 281 ----- 281  
 DB 5304 AATCCTTCTCGGGATACCCCTTCTTCTCTGCTCAGTTACCCCTTCCCTACTCCACAGA 5363  
 QY 282 -----ArgTrpGlyThrValCysHis----- 288  
 DB 5364 GCGCTCTCTCTCTGCGAGTACTCTGGGCGATATATTTTCAACCCCACTCTGTAAC 5423  
 QY 289 -----HisLysTrpAsnAlaAlaAla----- 296  
 DB 5424 GAGACCTTAGCATGGGGTCTTTAACACACATGGGATTTAGCTTCTGCTTCTTCTCAG 5483  
 QY 297 -----AspValValCys----- 300  
 DB 5484 TCCATCTCAGCTTTCATCAGAGTTCATGCTGTGCCAATGAACAACTCTTAGAGTTC 5543  
 QY 301 --LysGlnLeuGlyCys-----GlyThrAlaLeuHis-PheAlaGlyLeu----- 314  
 DB 5544 AGAAAGAGCCTGCTGTTTAAATGAGCTTTGGCCTCTTTATATTAGGCTGACACAACT 5603  
 QY 314 ----- 314  
 DB 5604 TTCGGGGAATTGAGGTTGAGACTCTCTAATGTTCTGTTGAAGGCAAGGTCACTACAGGCT 5663  
 QY 315 -----ProHisLeu----- 317  
 DB 5664 TGATACCCCATCTCTCACTACTTGAACACCCAGATTTTGCAGGCGCCAGCTCTGCAT 5723



Db	7840	TTGCTGGGGTGGAGTTCTTGACCTCAGCTCTTCTCAGAAGCGTGTGCAGCAATTGCCCTG	7899
QY	682	CysSerAspAlaSerAspMetGluLeuArgLeuValGlyGlySerSerArgCysAlaGly	701
Db	7900	TGTTCTAGGTCGTGGTAG-----GGG	7920
QY	702	LysValGluValAsnValGlnGlyAlaValGlyIleLeuCysAlaAsnGlyTrpGlyMet	721
Db	7921	AGGCGAGTCCCATGAGGCCAGCACATGGCCTTGTCATTCGCTGCTGATTGGGCGCTG	7980
QY	722	AsnIleAlaGlu-----ValValCysArgGlnLeuGluCysGlySerAla	736
Db	7981	AAGATCGCAACAAGGATTTTGGCTGGAGTGGCTTCTCAGCCITTCGTACTCAGGAACAC	8040
QY	737	IleArgValSerArgLiu-----ProHisPheThrGluArgThrLeu	750
Db	8041	CTAAGATGTGCRAAGGAGTGGGTGGTTAGGTCAACGGGTACCCTGGGCAGACACAA	8100
QY	751	-----HisIleLeuMetSerAsnSergLy-----CysThr-GlyGl	762
Db	8101	TTTGATCACTCAGAGCTGGCAATAGTGGCAGGATCTGCCTGCACCCCTTACACGGT--	8158
QY	762	yGluAlaSerLeuTrpAspCysIleArgTrpGluTrpLysGlnThrAlaCysHisLeuAs	782
Db	8159	---GCATCTGTGGGGATG---GCATGGCAATG---CCCTCCCTGTGTATAGGAA	8208
QY	782	nMetGlu-----Alase	786
Db	8209	CTAGGATGGACTGAGTGTCAAGACTCGCTCATTTCTTCCCTCCTCGTCCAGATTTTGC	8268
QY	786	rLeuIleCysSer-----	790
Db	8269	ACTTCTGTGTAATGTTCTCGATCTGACCTTCTCTCTTCTTCTCAGACTCCCACAGTCC	8328
QY	791	-AlaHisArgGlnProArgLeuValGlyAlaAspMetPro-----	803
Db	8329	GCGCGACCCAGCCAGGTAGTCCCGAGTGTCTCTCTCAAATGTCCCTTCTCTTC	8388
QY	803	-----	803
Db	8389	TGCCAATCACCCCTTCCACACTCCACAGAGCTCTCTCTCTCTGTGTGGATACTGTG	8448
QY	803	-----	803
Db	8449	GGGCATATTAATTTCCACCCCCAACACCGCTGTGTAACTGAGATCCAGACACAGCGCTT	8508
QY	803	-----	803
Db	8509	TTAAACACACACAGGATTGGAGGGCCCTGTGCTCTTTTTTCAACCCCTCTCAGCTTTTCAT	8568
QY	803	-----	803
Db	8569	GAACAGTTTTCATACTGTCCAGTGGACAAACCTTTACAGTTTCAGAGTGGCCCCCATGT	8628
QY	803	-----	803
Db	8629	TTAATGAGCTTTGGTCCTTTTATTTATTCGGAGTCAATATAGTTTCTGAAAATTGAGGGT-	8688
QY	804	-----CysSerglyArgValGluValLysHis-----	812
Db	8689	GTCACCTCTGTAACCGTTCAAGGCATCGTCAAGGCGAGCTCGATACCCCATCCATCACT	8748
QY	813	-AlaAspThrTrpArgSerValCysAspSerAspPheSerLeuHisAlaLaasnValLe	832
Db	8749	GCTGGAACAACTCCGAGATTAT---GATTCCGAGATTCTATCCAGAGAGGCGCAGATTT	8805
QY	832	uCys-----ArcGluLeu---AsnCysGlyAspAlaIleSerLeu-----	844
Db	8806	GTGCTCGGGCAGGAGGAGGATATAAAGTTT- GTGATGTCTCTGTCTATACAGAAACC	8864
QY	844	-----	844

Db	8865	TGATTCTCGATTGTCCTCCCTGGCGACGCCCTGGTTCCTCCCTAACATTTTACCTGGCAGTGTC	8922
Qy	845	-----SerValcIyAs	848
Db	8925	CGAGCTTCAGCAATGGCGCTCTCATGTCCTAGTCAGTCCATGCATCAGCAGATGTGGGATG	8984
Qy	848	PHISpHeGlyLysGlyAsnGlyLeu	856
Db	8985	GCAT-----GGTGGGGCATCTCTTAACAGATAGAAAGATACCCAGGATTACAGA	9035
Qy	856	-----	856
Db	9036	AATGGAGGGCTCAGTCTGGTCTCCAGCAGGACCTTTGTCCCTGGATGAGTTACACAGCACA	9095
Qy	857	---ThrTrpAlaGluLysPheGlnCysGluGlySerGluThrHisLeuAlaLeuCys-P	875
Db	9096	GGAGACCTGGGNAGACACATGGGAACAATGGCAGGAACAGAGTGGAAATGTTGTCTCA	9155
Qy	875	roIleValcIuNHISProGluAsp	882
Db	9156	CGTGTGATCTCCTCCACAGATCCTTTGTGTTCTGTGCTTTTCCCTTCAAGTCTAATTC	9215
Qy	883	-----ThrCysIleHisSera	888
Db	9216	TGTCCTTTCCCTTTGTGCAATTTACAGATACTTGGCCGACCTCACATGCATCAACAGCA	9275
Qy	888	rgGluVal-----GlyValValCysSerArgTyr	897
Db	9276	GGTAATAACCTCTCACCCCTCCCTAGGACTACATATCTCTGGACATATTTTGTGTTG	9335
Qy	898	---ThrAsp---ValArgLeuValAsnGlyLysSer	907
Db	9336	AAACTGATGATGAGGTCAATGTGGCTTCTCTGTTTTCATGTCCTCGGTGGGTTCGGT	9395
Qy	908	--GlnCysAspGlyGlnValGluIleAsnValLeuGly	919
Db	9396	GGGAGAGGTGGAAATCTGTAGAGGACCACTCTGGGTCTGATGTTTGAGGACGGAGGT	9455
Qy	920	-----HisTrpG	922
Db	9456	GTTGGTGACCTCTCCCATGGAATCCTGTTCCAAGTGGTCAGGAAGATCCTCATCCAG	9515
Qy	922	lySerLeuCysAspThrHisTrp	931
Db	9516	GTGCTCAG---CACGAGCACTGGAGGGCTCTTAATGCTGCTGGGACCTCATTCCTGGCC	9572
Qy	931	roGlu-----AspAlaArgValLeu	943
Db	9573	CTCAGGCCATGGGATCCAGACCTCTTAAGAGCGGGTGAAGTCCCGGTCCCTGCACCTGT	9632
Qy	943	ysGlyThrAlaLeuSerThrThrGlyGlyLysTyrIleGly-GluArgSerValArgVal	962
Db	9633	GTGGCCAGGCCCTTGCCATCATCGCAATTTGCCAAGGACAGAGGCCCATGTCAGGTG	9692
Qy	963	-----TrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeuLeu	976
Db	9693	CCAATAAGCTCTGAAATATGG-----AGGGGTCTAGGCCCTGTCATT	9734
Qy	977	AspAsnCysGlnMetThrValLeuGlyAlaProProCysIleHis	991
Db	9735	TGAGCTGTGTA-----GTCCATCTCTGTGTTCCACCAGAGTGGGAGTG	9779
Qy	992	-----GlyAsnThrValSerValIleCysThrGlySerLeuThr	1004
Db	9780	GGGTGCCATTCTCTCTCTGGGTGCATACATGCTTACCATACCTTGACGCTTT	9839
Qy	1005	-----GlnProLeuPheProCysLeuAlaAsnValSerAsp	1016
Db	9840	TATAGACTGTAGATAAGGCATCACATTTTCCCATCTATGACCAAGCTTAACCTCTGGG	9899
Qy	1016	-----	1016
Db	9900	TGCAGCCATCTGCCATTGTGACTGCGTGGCCTTGACACTTTGGCATTTTGATTTCAAACCT	9959

Qy 1017 ProTyrLeuSerAlaValProGlu----- 1024  
 Db 9960 GACFACCTTTCCCATGTCCTGAGTGTGTCATCCCTTTCCCTTCACCTCAGATGAG 10019  
 Qy 1025 ---GlySerAlaLeuLeuLeuCysLeu-----GluAspLysArgLeuArgLeu 1038  
 Db 10020 ATGATGAAGGATTTCTTGTGTTCCCTCTAGGACCTGAATCCAGTTTGGCCCTCAGGCTG 10079  
 Qy 1039 ValAspGlyAspSerArgCysAlaGlyArgValGluLeuTyrHisAspGlyPheTrpGly 1058  
 Db 10080 GTGAATGAGGTGACAGGTGTCAGGCGGAGTGGAGGTCCTATACCGAGGTCCTCGGGGC 10139  
 Qy 1059 ThrLeuCysAspAspGlyTrpAspLeuSerAspAlaHisValValCysGlnLysLeuGly 1078  
 Db 10140 ACCGTGTGTGATGATGCTGGACACACAGTCAGCGCATGTGCTGCGCGCAGTGGGC 10199  
 Qy 1079 CysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGluGlySerGlyProLeu 1098  
 Db 10200 TGTGGCTGGCCACGTCAGCCAGCAATGCCGGTTTGGCCAGGTTTCAGGACCCATT 10259  
 Qy 1099 TrpLeuAspAspLeuAsnCysThrGlyThrGluSerHisLeuTrpGlnCysProSerArg 1118  
 Db 10260 GTCTGTGATGAGTGCCTGCTCAGGCTATGATGCTCTACCTGTGGAGTGCCTCATATATCT 10319  
 Qy 1119 GlyTrpGlyGlnHisAspCysArgHisLysGluAspAlaGlyValIleCysSer----- 1136  
 Db 10320 GCGTGGCTCTCCATACTGTGACGACAGTGAAGACGCTGGTGTCTCTGCTCAGGTGG 10379  
 Qy 1136 ----- 1136  
 Db 10380 CCTCCAGACCTTGGGCTCCCTCTCTTGGGTAGATTTTGTCCAGAAAGCGAGGTCTCAT 10439  
 Qy 1137 -----GluPheThrAlaLeuArgLeuTyr----- 1144  
 Db 10440 TATGTTCTGATCTCTACTCAGAGCTTTTTCAGGCTTTCTATATATATCTGATATCTCT 10499  
 Qy 1145 -----SerGluThrGluThrGluSer-Cys----- 1152  
 Db 10500 TAGCTCTCTCTAGGAACATGCATGAGCTTCTCATTTGCCAGGTTTTCAGGAGTCCAGGTAG 10559  
 Qy 1153 -----AlaGlyArgLeuGluVal-----PheTyrAsnGlyThrTrpGln 1165  
 Db 10560 GACAACGGCCCAAGTGAATGAAGGTTCACGCTTTGTTCACCTACCGAGGACGAGC---GC 10616  
 Qy 1165 ySerValGlyArgArgAsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGln 1185  
 Db 10617 AAGCAGAGGAGAGAGAGAAAGTCCAGGCTCTTGGC-----TTTAA 10658  
 Qy 1185 yCysGlyGluAsnGlyValValSer-----LeuAlaProLeuSerLysThrGln 1201  
 Db 10659 GTGTGGCTGGAAGAAATAGTGGGCTGGTGTGTCATGAGAAAGAGGCTGATTTGGG 10718  
 Qy 1201 ySerGlyPheMetTrpValAspAspIleGlnCysProLysThrHisIleSerIleTrpGln 1221  
 Db 10719 TCTTGGCTTTGCA----- 10733  
 Qy 1221 nCysLeuSerAlaProTrpGluArgArgIleSerSer----- 1233  
 Db 10734 CTGCATA---GGACATGCTGACGCGCTCTCAGCTTCTGATGAGGCGCTCGGGCTGC 10790  
 Qy 1234 -----ProAlaGluGluThrTrp-----IleThrCysGluAspArgIleArgVa 1248  
 Db 10791 TTGTACTCTAGCAAGAGGATGGGAGGTGGCTGGCCTGGCAGAGGTCAACAAGGTGAGGGA 10850  
 Qy 1248 lArgGlyGlyAspThrGluCysSerGly----- 1257  
 Db 10851 GGAAGGAAGAGAGACCAAGGAGGAGGCTCAGACTCTTCAACCAACCACTCTTGTCTAGGA 10910  
 Qy 1257 ----- 1257  
 Db 10911 ACTAAGAAGAAGACTCACCCCTGACCGAGGAGGACCTAAGTTATTATCATGAGTGTGGT 10970

Qy 1258 ---ArgValGluIleTrp-----HisAlaGlySerTrpGlyThrValCysAs 1272  
 Db 10971 CTCCATGACCCAGACCTGGTGCATCAGGCTCACCTCTAAACTTGGGATTCAGTTGCCAA 11030  
 Qy 1272 pAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGlyCys----- 1289  
 Db 11031 TATAACA-----CTTGGATGTGACAAGCC 11054  
 Qy 1290 -----GlySerAlaLeuAlaAlaLeuArgAspAlaSer 1300  
 Db 11055 TCTAAAGTATAGCACCCCTGGCTCTCCAGGTTCCATCTTCCCTACCTAGCAAGAT----- 11109  
 Qy 1300 rPheGlyGlnGlyThrGlyThrIleTrpLeuAspAspMetArgCysLysGlyAsnGluSer 1320  
 Db 11110 -TTTCAAAAGGT-----TACATGATGAA----- 11133  
 Qy 1320 rPheLeuTrpAspCysHisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAs 1340  
 Db 11134 -----GCACATATATAGGA 11147  
 Qy 1340 pAlaGlyValArgCysSerGlyGlnSerLeuLysSerLeuAsnAlaSerSerGlyHisLe 1360  
 Db 11148 TGGGGA-----GGGACTCTCTCTCTGGGAGGAGCCCAAGAACACAG 11189  
 Qy 1360 uAlaLeuIleLeuSerSerIlePhe 1368  
 Db 11190 GCTCTCTGCTAGGAACGCGGCTTC 11214  
 RESULT 4  
 US-09-341-587-2  
 ; Sequence 2, Application US/09341587  
 ; Patent No. 6346606  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mollenhauer, Jan  
 ; TITLE OF INVENTION: Protein Containing an SRCR Domain  
 ; FILE REFERENCE: 4121-108  
 ; CURRENT APPLICATION NUMBER: US/09/341.587  
 ; CURRENT FILING DATE: 1999-08-31  
 ; EARLIER APPLICATION NUMBER: PCT/DE98/00096  
 ; EARLIER FILING DATE: 1998-01-09  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 2  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-341-587-2  
 Alignment Scores:  
 Pred. No.: 1,51e-56 Length: 2001  
 Score: 715.50 Matches: 164  
 Percent Similarity: 48.18% Conservative: 61  
 Best Local Similarity: 35.12% Mismatches: 119  
 Query Match: 8.79% Indels: 123  
 DB: 4 Gaps: 15  
 US-09-759-130B-381 (1-1453) x US-09-341-587-2 (1-2001)  
 Qy 956 GlyGluArgSerValArgValTrpGlyHisArgPheHisCysLeuGlyAsnGluSerLeu 975  
 Db 1 GGACACGAGTCTTANAGCTGTGAGC-----TGCCCCCACTAGCTGGCTC 48  
 Qy 976 LeuAspAsnCysGlnMetThrValLeuGlyAlaProCysIleHisGlyAsnThrVal 995  
 Db 49 TCCCAACACTGGGC-----CATCATGAAGATGCT 78  
 Qy 996 SerValIleCysThrGlySerLeuThrGln-----ProLeuPheProCysLeuAlaAsn 1013  
 Db 79 GGTGTCTATCTGCTAGCTGCTCAGTCCAGTCAAGCCCGCCAGCATACTTGGCTGACC 138  
 Qy 1014 ValSerAspProTyrLeuSerAlaValProGluGlySerAlaLeuIleCysLeuGluAsp 1033  
 Db 139 ACCAACTTACCGGCATTGACAGTAGGATCTGAATCCAGTTTG----- 180

QY 1034 LysArgLeuArgLeuValAspGlyAspSerArgCysAlaGlyArgValGluIleTyrHis 1053  
|||||  
Db 181 ---GCTCTGAGGTGGTGAATGAGGTGACAGGTGTCGAGGCGGAGTGGAGTCTGTAT 237  
QY 1054 AspGlyPheTrpGlyThrIleCysAspGlyTyrAspLeuSerAspAlaHisValVal 1073  
|||||  
Db 238 CGAGGCTCTGGGACCGCTGTGTGATGACAGCTGGGACCAATGATGCCAATGTGGTC 297  
QY 1074 CysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHisPheGlyGlu 1093  
|||||  
Db 298 TCGAGCAGCTGGGTGTGGCTGGCCATGTGCGCCAGGAATCCCGGTTGGCCAG 357  
QY 1094 GlySerGlyProIleTrpLeuAspLeuAsnGlyThrGlyThrGluSerHisLeuTrp 1113  
|||||  
Db 358 GGCTCAGGACCATTTCTCTGATGATGTGCTGCTCAGGAAATGAGTCTACCTGTGG 417  
QY 1114 GlnCysProSerArgGlyTrpGlnHisAspCysArgHisLysGluAspAlaGlyVal 1133  
|||||  
Db 418 AGCTGCCCCACAAAGCTGGCTCACCAACATGTGGCCATCAGGAAGCTGGTGTTC 477  
QY 1134 IleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThr----- 1147  
|||||  
Db 478 ATCTGCTCA-----GCCACCAATAAATTTCTACTACGACATGGTGGCATCCA 528  
QY 1148 -----GluThrGluSerCysAlaGlyArgLeuGluValPheTyr 1160  
|||||  
Db 529 ACAACTACAACTGCAAGACCTCTTCAAAATGTGGTGGCTCTTA-----TTCTAT 582  
QY 1161 ---AsnGlyThrTrpGlySer----- 1166  
|||||  
Db 583 GCCAGTGGGACATTTCTCCAGCCCATCTACCTGCTACATACCCCAACAATGTAAGTGT 642  
QY 1167 -----ValGlyArgArgAsnIle----- 1172  
|||||  
Db 643 GTTGGGAATAGAGTGAATTTCTGGTTATCGCATAAACCTGGCTTCAGTAACTCGAAA 702  
QY 1172 ----- 1172  
Db 703 TTGGAGGCACACATAACTGTCAGTTTGTATTGTTGAAATCTTTGATGGATCATTTGAAT 762  
QY 1173 ThrThrAlaIleAlaGlyIleValCys-----ArgGlnLeuGlyCysGlyGluAsn 1189  
|||||  
Db 763 ACAGTCTCTCTGGGAAATCTGTATGATACAGCAATATTTACATCTCTTAC 822  
QY 1190 GlyValValSerLeu-----AlaProLeuSerLysThrGlySerGlyPheMet----- 1205  
|||||  
Db 823 AACCGAATCACCATTACATTTTGAAGTGACATCAGTTTCCAAACACATGGCTTTTGGCT 882  
QY 1206 TrpValAspPheIleGlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAla 1225  
|||||  
Db 883 TGGTATAAC----- 891  
QY 1226 ProTrpGluArgIleSerSerProAlaGluGluThrTrpIleThrCysGluAspArg 1245  
|||||  
Db 892 -----TCTTCCCAAGCATGCCACCTTG-----AGGTG 921  
QY 1246 IleArgValArgGlyCysAspThrGluCysSerGlyArgValGluIleTrpHisAlaGly 1265  
|||||  
Db 922 GTCAATTTAAATTCATCTCTATGCTGTGCGGGGCGGTGAGAAATTTACCATGGTGC 981  
QY 1265 SerTrpGlyThrValCysAspSerTrpAspLeuAlaGluValValCysGln 1285  
|||||  
Db 982 ACCTGGGGACAGTTTGTGATGACTCTTGACCATTTACGAAAGCTGAGGTGCTGCAGA 1041  
QY 1286 GlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGlyThr 1305  
|||||  
Db 1042 CAGCTAGGTGTGGAGCTGCATTTTCAGCCCTTGAATATGTCATATTTGGCTCTGCTCT 1101  
QY 1306 GlyThrIleTrpLeuAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCys 1325  
|||||  
Db 1102 GGCCCCATCACCTGGACCATAGAGTCTCAGGAGCGGAATCCACTCTCTGGCAGTGC 1161

QY 1326 HisAlaLysProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCys 1345  
|||||  
Db 1162 CGGAACCGAGGTGGTCTTCCCAACAACATGTAATCATCTGTAAGATGCTGTGTCATCTGC 1221  
QY 1346 SerGlyGlnSerLeuLysSer 1352  
|||||  
Db 1222 TCAGGAACCATCTATCGACA 1242  
RESULT 5  
US-09-034-916-1  
; Sequence 1, Application US/09034916  
; Patent No. 6046314  
; GENERAL INFORMATION:  
; APPLICANT: GEBE, JOHN A.  
; APPLICANT: SIADAK, ANTHONY W.  
; APPLICANT: ARUFFO, ALEJANDRO A.  
; TITLE OF INVENTION: SPALPHA: A NOVEL SCAVENGER RECEPTOR  
; TITLE OF INVENTION: CYSTEINE-RICH DOMAIN-CONTAINING POLYPEPTIDE AND MONOCLONAL  
; TITLE OF INVENTION: ANTIBODIES THEREOF  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBINS & ASSOCIATES  
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
; CITY: MENLO PARK  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,916  
; FILING DATE: 04-MAR-1998  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/039,956  
; FILING DATE: 06-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5998-0020  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 325-7812  
; TELEFAX: (650) 325-7823  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2178 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 64..1101  
US-09-034-916-1  
Alignment Scores:  
Pred. No.: 2,94e-54 Length: 2178  
Score: 691.50 Matches: 148  
Percent Similarity: 52.51% Conservative: 40  
Best Local Similarity: 41.34% Mismatches: 139  
Query Match: 8.50% Indels: 31  
DB: 3 Gaps: 9  
US-09-759-130B-381 (1-1453) x US-09-034-916-1 (1-2178)  
QY 21 LeuPheSerAlaValValThrCysIleLeuLeuAsnSerCys-----Phe 36  
|||||  
Db 70 CTATCTCC-----TTGATCCCTTGCATTTGCACAGACCTGGATTC 111  
QY 37 LeuIleSerSerPheAsnGlyThrAspLeuGluLeuArgLeuValAsnGlyAspGlyPro 56

```
Db 112 CTAGCGTCTCCATCTGGA-----GTGCGGCTGGTGGGGGCTCCACCGC 156
Qy 57 CysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAspGly 76
Db 157 TGTGAAGGCGGGTGGAGGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 216
Qy 77 TrpAsnThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSerPheAla 96
Db 217 TGGACATTAAGGACGTGGCTGTGTGGCGGAGCTGGGCTGTGGAGTGGCCAGCGGA 276
Qy 97 Met-----PheArgPheGlyGlnAlaValThrArgHisGlyLysIleTrpLeuAsp 113
Db 277 ACCCTAGTGTATTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 336
Qy 114 AspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHisArgGluTrpGly 133
Db 337 TCAGTCAAGTTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 390
Qy 134 SerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
Db 391 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 450
Qy 148 TyrGlyGluAlaAsnLeuGluValArgLeuValAspGlyAsnAsnSerCysSerGlyArg 167
Db 451 TTCCTCCAGTCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 510
Qy 168 ValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsn 187
Db 511 GTGAAGTGAAGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 570
Qy 188 ThrAlaAlaValValCysArgGlnLeuGlyCysProSerSerPheIleSerSerGlyVal 207
Db 571 GCCGCAAGAGTGGTGGCGGACGCTGGATGTTGGGAGGCTGTACTGACTCAAAAGC 630
Qy 208 ValAsnSerProAlaValLeuArg---ProIleTrpLeuAspAspIleLeuCysGlnGly 226
Db 631 TGCAACAAGCATGCTGTGGCGGAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGA 690
Qy 227 AsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAspCysSerHis 246
Db 691 CGAAGAACCAACCTTCAGGATGGGCTTGGGCTTGGGGAAGAACACCTGCAACCAT 750
Qy 247 AsnGluAspValThrLeuThrCysTyrAspSerSerAspLeuGluLeuValGly 266
Db 751 GATGAAGACACGTGGTGGTGAATGTAAGATCCCTTTGAC-----TTGAGACTAGTAGGA 804
Qy 267 GlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrpGlyThrVal 286
Db 805 GGAGACAACCTCTGCTCTGGGCGACTGGAGGTGCTGCAACAAGGCGGTATGGGGCTCTGTC 864
Qy 287 CysHisHisLysTrpAsnAsnAlaAlaAspValValCysLysGlnLeuGlyCysGly 306
Db 865 TGTGATGACACTGGGGAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 924
Qy 307 ThrAlaLeu-----HisPheAlaGlyLeuProHisLeuGlnSerGlySerAspValVal 324
Db 925 AAGTCCCTCTCTCCCTCTCCAGAGACCGGAAATGCTATGCGCTGGGTTGGCGGCATC 984
Qy 325 TrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCysArgHisSer 344
Db 985 TGGCTGGATAATGTGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1044
Qy 345 GlyThrValAsnPheAspCysLeuHisGlnAsnAspValSerValIleCysSer 362
Db 1045 TTTTGGGGGTTTACGACTGTCACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098
```

## RESULT 6

US-09-276-400-3

; Sequence 3, Application US/09276400

; Patent No. 6140056

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran

```

; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND
; FILE OF INVENTION: USES THEREFOR
; FILE REFERENCE: MNI-073
; CURRENT APPLICATION NUMBER: US/09/276.400
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2262
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-276-400-3

Alignment Scores:
Pred. No.: 6,12e-49 Length: 2262
Score: 634.50 Matches: 214
Percent Similarity: 40.03% Conservative: 97
Best Local Similarity: 27.54% Mismatches: 253
Query Match: 7.80% Indels: 215
DB: 3 Gaps: 32
```

US-09-759-130B-381 (1-1453) x US-09-276-400-3 (1-2262)

```
Qy 30 LeuLeuLeuAsnSerCysPheLeuLeuSerSerPheAsnGlyThr----- 44
Db 40 CTGCTGCTG-----TGCCTGCTGTGTGCAAGTTCGTGCTTGGGGTCTCCCTTCCACG 93
Qy 45 -----AspLeuGluLeuArgLeuValAsn---GlyAsp 54
Db 94 GGCCTTGAGAGAGAGCGGCGGAGCGGCTTCGGTTCGCGCTGGCTTCCCGAGG 153
Qy 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db 154 AAGCCTACGAGGCGCGCTGGAGATACAGCAGAGTGGTGAATGGCGCACCATCTGCGAT 213
Qy 75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuGlyCysProPheSer 94
Db 214 CATGACTTTCAGCTGTCAGGCTGCCACATCTCTCTCCGGAGCTGGGC-----TTCACA 267
Qy 95 PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
Db 268 GAGGCCACAGGCTGACCCACAGTCCCAATATGGCCCTGGGAACA-----GGCCGC 318
Qy 110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpLeuCysGlnHis 129
Db 319 ATCTGGCTGGACAATTCAGCTGCAGTGGGACCGAGAGAGTGTGACTGAATGCTCC 378
Qy 130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
Db 379 CGGGCTGGGGGAACAGTGAAGTTCAGCGCAGATGAGGATGCTGGGGTCTATCTGCAAGAC 438
Qy 148 -----TyrGlyGluAlaAsnLeu----- 153
Db 439 CAGCCCTCCCTGGCTTCTCGGACTCCAATGCTATTGAGGTAGAGCATCACCTGCAAGTG 498
Qy 154 -----GlyLeuArgLeuValAspGlyAsnAsn 162
Db 499 CAGGAGGTGCGAATTCGACCCCGCTGGGTGGGCGAGACACACCTCGCC----- 549
Qy 163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAsp 182
Db 550 CTGACGGAGGGCTGGTGAAGTTCAGGCTTCTGACGGCTGGTTCGCAAGTGTGCGACAAA 609
Qy 183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201
Db 610 GCGTGGAGCGCCACACAGCAGCCAGCTGGTCTCGGGAGTGTGGCTTCCCGAGGAAAG 669
Qy 202 PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le 219
Db 670 AGGGTCAACGC---GGCTTCTACAGGCTGCTAGCCCCACCGCAGCAACATCTCTTGGTCT 728
Qy 219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
```

Qy	550	erAsnIleTrp----	AspCysGluHuHisSerGlyTrpGlyHisAsnCysValHisArgG	569
Db	1799	CCTGGGTGGCAGAGTGC-----	-----CATGGCATTACCACAGCA	1837
Qy	569	luaspVal-----	-----IleValThrCysSerGlyAspAlaThrTrpGlyLeuA	584
Db	1938	TGGACATCTTCACACTACTATGATATCTCACC	CCCAATGGCCACCAAGTG-----	1887
Qy	584	rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTrpPheGlnGlyArgT	604	
Db	1888	-----	-----GCTGAGGCGCCACA	1900
Qy	604	rpGlyThrValCysAspAspGlyTrpAsnSerIysAlaAlaValValCysSerGlnL	624	
Db	1901	AAGCTAGTTCTCTG-----	-----	1914
Qy	624	euaspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL	644	
Db	1914	-----	-----	1914
Qy	644	ysIleTrpLeuAspValSerCysAspGlyAspGlySerAspLeuTrpSerCysArgA	664	
Db	1915	-----CTCGAAGACACTGAGTGTGAGGAGGATGCTCTCCAAGCGGTATGAGTGTGCCA	1966	
Qy	664	snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA	684	
Db	1967	ACTTTGGA-----	-----GAGCAGGCGTACTGTGGGTTGCTGGG	2020
Qy	684	sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg	698	
Db	2003	ATCTCTACCGGCATGACATTGACTGTGCTGAGTGGATTGACATCAGG	2047	

RESULT 7

US-09-448-076-3

; Sequence 3, Application US/09448076  
; Patent No. 6300092  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran et al.  
; TITLE OF INVENTION: METHODS OF USE OF A NOVEL LYSYL OXIDASE-RELATED PROTEIN  
; FILE REFERENCE: MNI-073CP  
; CURRENT APPLICATION NUMBER: US/09/448,076  
; CURRENT FILING DATE: 1999-11-23  
; EARLIER APPLICATION NUMBER: 60/117,580  
; EARLIER FILING DATE: 1999-01-27  
; EARLIER APPLICATION NUMBER: 09/276,400  
; EARLIER FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2262  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-448-076-3

Alignment Scores:

Pred. No.: 6.12e-49 Length: 2262  
Score: 634.50 Matches: 214  
Percent Similarity: 40.03% Conservative: 97  
Best Local Similarity: 27.54% Mismatches: 253  
Query Match: 7.80% Indels: 215  
DB: Gaps: 32

US-09-759-130B-381 (1-1453) x US-09-448-076-3 (1-2262)

Qy

30

LeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGlyThr-----

44

Db

40

CTGCTGCTG-----TGCCTGCTGCTCAGTTCGTCTGGGGTCTCCGTCCTCCCTCCACG

93

Qy

45

-----AspLeuGluLeuArgLeuValAsn---GlyAsp

54

Db

94

GGCCCTGAGAAGAAGCCGGAGCCAGGCGCTTCGGTTCCGCTGGCTTCCCCAGG

153



Db 2003 ATCTACGGCATGACATGACTGCTCAGTGGATTGACATCACGG 2047

## RESULT 8

US-09-702-572-3

; Sequence 3, Application US/09702572

; Patent No. 6391602

; GENERAL INFORMATION:

; APPLICANT: Khodadoust, Mehran

; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND

; FILE REFERENCE: MNI-073

; CURRENT APPLICATION NUMBER: US/09/702,572

; CURRENT FILING DATE: 2000-10-31

; PRIOR APPLICATION NUMBER: 09/276,400

; PRIOR FILING DATE: 1999-03-25

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 2262

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-702-572-3

## Alignment Scores:

Pred. No.:	6,12e-49	Length:	2262
Score:	634.50	Matches:	214
Percent Similarity:	40.03%	Conservative:	97
Best Local Similarity:	27.54%	Mismatches:	253
Query Match:	7.80%	Indels:	215
DB:	4	Gaps:	32

US-09-759-130b-381 (1-1453) x US-09-702-572-3 (1-2262)

QY	30	LeuLeuAsnSerCysPheLeuLeuSerPheAsnGlyThr-----	44
Db	40	CTGCTGCTG-----TGCCTGTGTGCACTTCTGCTTGGGTCTCCGCTCCCTCCACG	93
QY	45	-----AspLeuGluLeuArgLeuValAsn---GlyAsp	54
Db	94	GGCCCTGAGAAGAGCCGGAGCCAGGGCTTCGGTTCCTCGCTGGCTTCCCCAGG	153
QY	55	GlyProCysSerGlyThrValGluValLysPheGlnGlyThrPglThrValCysAsp	74
Db	154	AAGCCCTACAGGGCCGCGTGAGATACAGCAGCTGGTGAATGGGACCATCTCGCAT	213
QY	75	AspGlyTrpAsnThrAlaSerThrValCysLysGlnLeuGlyCysProPheSer	94
Db	214	GATGACTTCAGCGTGCAGGCTGCCCATCTCTGCCGGAGCTGGGC-----TTCACA	267
QY	95	PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys	109
Db	268	GAGGCCACAGCGCTGGACCCACAGTGCCTCAATATGGCCCTTGGAAACA-----	318
QY	110	IleTrpLeuAspValSerCysTyrglyAsnGluSerAlaLeuTrpGluCysGlnHis	129
Db	319	ATCTGCTGACAACTTGAGCTGCACTGGGACCGACAGAGTGTGACTGAATGTGCTCC	378
QY	130	ArgGluTrpGlySerHisAsnCysTyrglyHisGlyGluAspValGlyValAsnCys	147
Db	379	CGGGGTGGGGGAACAGTACTGTACCCAGCATGAGGATGCTGGGTCTCTGCAAAAGAC	438
QY	148	-----TyrGlyGluAlaAsnLeu-----	153
Db	439	CAGCGCTCCCTGGCTTCTCGGACTCCAATGTCATTGAGTAGAGCATCCCTGCAAGTG	498
QY	154	-----GlyLeuArgLeuValAspGlyAsnAsn	162
Db	499	GAGGAGTGGGAATTCGACCCCGCTGGTGGGGGAGAGACCCCTGGCC-----	549
QY	163	SerCysSerGlyArgValGluValLysPheGlnGlyThrPglThrIleCysAspAsp	182
Db	550	GTACGAGGGGCTGGTGAAGTTCAGGCTTCCTGACGGCTGGTTCGCAAGTGTGCGACAAA	609

QY	183	GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer-----Ser	201
Db	610	GGTGGAGCGCCCAACACACACAGCTGCTGCGGATGCTGGGCTTCCCCAGGAAAAG	669
QY	202	PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le	219
Db	670	AGGTCACAGC-GGCCTTCTACAGGCTCTAGCCCAACGGCAGCAACTCTCTTGGTCT	728
QY	219	uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr	239
Db	729	GCATGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT	788
QY	239	pCysAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys-----	254
Db	789	TCGTGCCCAATGACACCGCCAGGTGCCCTGGGGGGGGCCCTGCAGTGGTGTGTGCC	848
QY	255	-----TyrAspSerSer-----	258
Db	849	AGGCCCTGTCTACGGCGCATCCAGTGGCCAGAAAGCAACACAGTCTGAAGCCTCAGGG	908
QY	259	-AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy	278
Db	909	GGAGCCCGCTGTCTAAAGGGCGCGCCACCCCTGGAGAGGCGCGGTAGAAGTCTCT	968
QY	278	sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspVa	298
Db	969	GAAGCCAGCACATGGGCGACAGTCTGTACCGCAAGTGGGACCTGCATGCAGCAGCT	1028
QY	298	lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl	318
Db	1029	GGTGTCTGGGAGCTGGGCTTCGGGAGTCTCGAGAAGCTCTGAGTGGCGCTCGCATGG	1088
QY	318	nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe	338
Db	1089	CGAGGCGATGGTGTATCCACCTGAGTGAAGTTCGCTCTGGCAGAGAGCTCTCCCT	1148
QY	338	uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe	358
Db	1149	CTGGAAGTCCCGCCCAAGAACATCACAGCTGAGGATTGTCACATGACGAGGAGCGCGG	1208
QY	358	rValIleCysSer-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe	375
Db	1209	GTCTCGGTGCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT	1268
QY	375	rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----GlnTrpTr	391
Db	1269	CAGCCCAACATGAGGGCGAGTGCAGGTGCAATAGGGGACCTGGGCGCCCTTCGCTGGGG	1328
QY	391	pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl	411
Db	1329	CCTCATCTGTGGGTGATGCTGGGACCTGGGACCTGGGACCTGGGACCTGGGACCTGGG	1388
QY	411	yCysProPheSerValPheGlySerArgArgAla-----LysProSerAsnGluAlaAr	429
Db	1389	TCTGGCTACGCCAACACCGGCTGCAGGAGAGCTTGTACTGGGACTCTGGGAATAAATAC	1448
QY	429	gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh	449
Db	1449	AGAGTGGTGTGATGAGTGGAGTGCCTGTCACAGGAGTGCCTGCTGCTGCTGCTGCTG	1508
QY	449	rTyrrAspGlyLysAlaLysArgThrCysPheArg-----ArgSerAspAlaGlyVa	466
Db	1509	CATCATATGGC-----ACCCATCATCTGCAAGAGGACGAGGACCCGCTTCACTGCTGAGT	1565
QY	466	lIleCysSerAspLysAlaAspLeuArgLeuValGlyAlaHisSerProCysTy	486
Db	1566	CATCTGTTCTGAGACTGCATCA---GATCTGTGCTG-----CACTCA-----	1605
QY	486	rglyArgLeuGluValLysTyrrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe	506
Db	1606	-----GCATGGTGTGCA---GGAGACCGCCCTACATA	1630
QY	506	rThrArgAsnAla-----AlaValValCysLysGlnLeuGlyCysGlyLysProMe	523

US-09-759-130B-381 (1-1453) x US-09-276-400-1 (1-2920)

Qy	30	LeuLeuLeuAsnSerCysPheLeuLeuSerSerPheAsnGlyThr-----	44
Db	182	CTGCTGCTG-----TGCCTGTGTGCAGTTCTGTCTTGGGCTCCGCTCCCTTCCACG	235
Qy	45	-----AspLeuGluLeuArgLeuValAsn---GlyAsp	54
Db	236	GGCCCTGAGAAGAGCCGGAGCCAGGGGCTTCGGTTCGGCTGGCTGGCTTCCCGCAGG	295
Qy	55	GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp	74
Db	296	AAGCCCTACGAGGCGCGTGGAGATACAGCAGCTGTGAATGGGCGACCATCTCGCAT	355
Qy	75	AspGlyTrpAsnThrAlaSerThrValValCysLysGlnLeuGlyCysPropPheSer	94
Db	356	GATGACTTCAGCTGCAGGCTGCCACATCCTCTCCGGGAGCTGGC-----TTCA	409
Qy	95	PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys	109
Db	410	GAGGCCACAGCTGGACCCACAGTCGCAATATGGCCCTGGAACA-----GGCGC	460
Qy	110	IleTrpLeuAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis	129
Db	461	ATCTGGCTGACAACTTGAGCTGCAGTGGGACCGGACGAGAGTGTGACTGAATGT	520
Qy	130	ArgGlnTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys-----	147
Db	521	CGGGCTGGGGAAACAGTACTGTACGCACGATGAGGATGTGGGGTCACTCTGCAAGAC	580
Qy	148	-----TyrGlyGluAlaAsnLeu-----	153
Db	581	CAGGCGCTCCCTGGCTCTCGGACTCCAAATGTCAATGAGGTAGACATCACCTGCAAGTG	640
Qy	154	-----GlyLeuArgLeuValAspGlyAsnAsn	162
Db	641	GAGGAGGTGCGAATTCGACCGCGCTGGGTGGGCGACGACCCCTCC-----	691
Qy	163	SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp	182
Db	692	GTGACGAGGGGCTGTGGNAAGTCAGGCTTCCTGACGGCTGTGCCAAGTGTGGACAAA	751
Qy	183	GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser	201
Db	752	GGCTGGAGCGCCACAAACAGCCACGCTGTCTGGGGATGCTGGGCTTCCCGCAGCGAAAG	811
Qy	202	PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le	219
Db	812	AGGCTCAACGC-GGCTCTTACAGGCTGTGATCCCAACGCGCAGCAACACTCTTTGGTCT	870
Qy	219	uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr	239
Db	871	GCATGGGTGGCTGGCTGGGACGAGGCCACCTCTCCCTCTCTTCCCTGGAGTCTTA	930
Qy	239	pGlyAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys-----	254
Db	931	TCGTGCCAATGACACCGCCAGGTGCCCTGGGGGGGCCCTGCGAGTGGTGTGTGTGCC	990
Qy	255	-----TyrAspSerSer-----	258
Db	991	AGGCCCTGTCTAGCGCGCATCCAGTGGCCAGAGAACAAACACTGCGAAGCCTCAGG	1050
Qy	259	-AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy	278
Db	1051	GGAGGCCGTGTCCGTCTAAAGGCGGGCGGCCACCCCTGGAGAGGGCCGGGTAGAAGTCT	1110
Qy	278	sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspVa	298
Db	1111	GAAGCCAGGCACATGGGGCACAGTCTGTACCCCAAGTGGGACCTGCATGCGAGCCGCT	1170
Qy	298	lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl	318
Db	1171	GCTGTCTGGGAGCTGGGCTTCGGGAGTCTCGAGAAGCTCTGAGTGGCGGTTCGATGGG	1230

Qy	278	s	l	e	e	l	e	l	n	g	l	y	a	r	g	t	r	p	c	l	y	t	h	r	v	a	i	c	y	s	h	i	s	t	r	p	a	n	a	n	a	l	a	a	a	s	p	v	a	298			
Db	1111	G	A	G	C	C	A	G	C	A	T	G	T	G	A	C	C	A	G	T	G	G	A	C	C	T	G	T	G	A	C	C	A	G	T	G	G	A	C	C	A	G	C	G	T	1170							
Qy	298	l	v	a	i	c	y	s	l	e	u	d	i	c	y	s	g	l	y	t	h	r	a	l	a	l	e	u	i	s	p	h	e	a	l	a	g	i	y	l	e	u	p	r	o	h	i	s	l	e	u	1	318
Db	1171	G	E	T	G	T	C	G	G	A	G	T	C	G	G	A	G	T	C	G	A	A	A	G	T	C	T	G	A	G	T	C	G	A	A	G	T	C	T	G	A	G	T	G	G	C	T	G	C	A	T	1230	



Db 521 CGGGGCTGGGGAAACAGTACTAGCCAGCATGAGGATGCTGGGTCTCATCTGCAAGAC 580  
Qy 148 -----TyrGlyGluAlaAsnLeu----- 153  
Db 581 CAGGGCTCCCTGGCTTCGGGACTCCAATGTCTATTGAGGTAGAGATCACCTGCAAGTG 640  
Qy 154 -----GlyLeuArgLeuValAspGlyAsnAsn 162  
Db 641 GAGGAGTGGGAATTCGACCCCGCTGGGTGGGCGAGACGCCCTTGCCC----- 691  
Qy 163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182  
Db 692 GTGACGAGGGCTGGTGGGAAGTCAGGCTTCCTCAGCGGCTGGTGGCAAGTGTGCGACAAA 751  
Qy 183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuGlyCysProSer---Ser 201  
Db 752 GGTGGAGGCCCCACACACAGCCAGTGGTCTCGGGATGTCTGGGCTTCGCCACCAAGAAAG 811  
Qy 202 PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le 219  
Db 812 AGGTCAACGC-GGCCTTCTACAGGTGTGTAGCCCAACGCGCAGCAACACTCCTTTGGTCT 870  
Qy 219 uAspAspIleLeuGlyCysGlnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239  
Db 871 GCATGGGTGGCTGGCTGGGCGACGAGGCCACCTCTCCCTCTGTCTCCCTGGAGTTCTA 930  
Qy 239 pGlyAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys----- 254  
Db 931 TCGTGCACANTGACACCCGACAGTGCCTGGGGGGGCCCTGCAGTGTGAGCTGTGTGCC 990  
Qy 255 -----TyrAspSerSer----- 258  
Db 991 AGGCCCTGTCTACGCGCATCCAGTGGCCAGAGAACAAACAGTCGAGCCTCAGGG 1050  
Qy 259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278  
Db 1051 GGAGGCCGCTGTCCTGCTAAAGGGCGGCCGCCACCCCTGGAGAGGGCGGGTAGAAGCTCT 1110  
Qy 278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAlaAlaAlaAspVa 298  
Db 1111 GAAGGCAGCACATGGGGGACAGTCTGTGACCGCAAGTGGGACCTGCATGCAGCCACGCT 1170  
Qy 298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGl 318  
Db 1171 GGTGTGTCGGGAGCTGGCTTCGGGAGTCTCGAGAGAGCTGTAGTGGCGCTCGCATGGG 1230  
Qy 318 nSerGlySerAspValThrLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338  
Db 1231 GCAGGGCATGGGTGCTATCCACCTGAGTGAAGTTCGCTGTCTGGACAGGAGCTCTCCCT 1290  
Qy 338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358  
Db 1291 CTGGAAGTGGCCCAAGAACATACAGCTGAGGATGTTTCATAGCCAGGATGCGCG 1350  
Qy 358 rValIleCysSer-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspGlySe 375  
Db 1351 GGTCCGGTGAACCTACCTTACAGTGGGCGAGAGACCCAGGATCCGACTCAGTGGGGCGG 1410  
Qy 375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----GlnTrpTr 391  
Db 1411 CAGCCAAATAGGGGGGAGTCGAGGTGCAAAATAGGGGGAGCTGGGCCCTTCGCTGGG 1470  
Qy 391 pThrIleCysAspGlnAsnTrpLysAsnGluGlnAlaLeuValValCysLysGlnLeuGl 411  
Db 1471 CCTCATCTGTGGGATGACGTGGGGGACCCCTGGAGCCCATGGTGGCTGTAGGCAACTGGG 1530  
Qy 411 yCysProPheSerValPheGlySerArgArgAla-----LysProSerAsnGluAlaAr 429  
Db 1531 TCTGGCTACGCCCAACACCGGCTTCAGGAGACCTGTGTACTGGGACTCTGGGAATATAAC 1590  
Qy 429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAspCysTh 449  
Db 1591 AGAGTGTGTAGTGTGAGTGCCTGCACAGGAGTGCAGTGTCCCTGGATCATGCTGC 1650

Qy 449 rTyrAspGlyLysAlaLysArgThrCysPheArg-----ArgSerAspAlaGlyVa 466  
Db 1651 CCATCATGGC---ACCACATCCTGCAAGAGGACAGGACCGCTTCTACTGCTGGAGT 1707  
Qy 466 lIleCysSerAspLysAlaAspLeuAspLeuArgLeuValGlyAlaHisSerProCysTy 486  
Db 1708 CATCTGTCTGAGACTGCATCA---GATCTGTGTGCTG-----CACTCA----- 1747  
Qy 486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506  
Db 1748 -----GCACGTGGTGA-GGAGACCCCTTACA 1772  
Qy 506 rThrArgAsnAla-----AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523  
Db 1773 TCGAAGACCGGCCCTCATATGTTGTACTGTGC-----TGGGAAGAGAAT 1820  
Qy 523 t-----HisValPheGlyMetThrTyrPheLysGluAlaSerGlyProIl 538  
Db 1821 GCCTGGCCAGCTCAGCCCGCTCAGCCAACTGGCCCTATGTTCAACCGCTGTGCTCCGAT 1880  
Qy 538 eTrpLeuAspAspVal-SerCys-----IleGlyAsnGluS 550  
Db 1881 TCCTCTCCAGATCCACACACTGGGAGAGCTGACTTCAGCCCAAGCTGGGCGCCACT 1940  
Qy 550 eAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArg 569  
Db 1941 CCTGGGTGTGGCAGAGTGC-----CATGGCATTACCACAGCA 1979  
Qy 569 luAspVal-----IleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584  
Db 1980 TGGACATCTTCTACTCATATGATCTCACCACCAATGGCACCAGGTG----- 2029  
Qy 584 rGluValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyrPheGlnGlyArgT 604  
Db 2030 -----GCTGAGGGCCACA 2042  
Qy 604 rPglyThrValCysAspAspGlyTrpAsnSerLysAlaAlaAlaValValCysSerGlnL 624  
Db 2043 AAGTACTGTTCTGT----- 2056  
Qy 624 euAspCysProSerSerIleIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrGlyL 644  
Db 2056 ----- 2056  
Qy 644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArg 664  
Db 2057 -----CTCGAAGACACTGAGTGTGAGGAGGATGCTCCAAGCGGTATGAGTGGCCA 2108  
Qy 664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684  
Db 2109 ACTTTGGA-----GAGCAAGGCATCATCTGGTGTGCTGGG 2144  
Qy 684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698  
Db 2145 ATCTCTACCGCATGACATTGACTGCTCAGTGGATGATGACATCACC 2189

## RESULT 11

US-09-702-572-1  
; Sequence 1, Application US/09702572  
; Patent No. 6391602  
; GENERAL INFORMATION:  
; APPLICANT: Khodadoust, Mehran  
; TITLE OF INVENTION: NOVEL MSP-18 PROTEIN AND NUCLEIC ACID MOLECULES AND  
; TITLE OF INVENTION: USES THEREFOR  
; FILE REFERENCE: MNI-073  
; CURRENT APPLICATION NUMBER: US/09702, 572  
; CURRENT FILING DATE: 2000-10-31  
; PRIOR APPLICATION NUMBER: 09/276, 400  
; PRIOR FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1

```

; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (143)..(2401)
US-09-702-572-1

Alignment Scores:
Pred. No.:          Length:          2920
Score:             Matches:          214
Percent Similarity: 40.03%
Best local Similarity: 27.34%
Query Match:       Indels:           253
DB:                Gaps:             32

US-09-759-130B-381 (1-1453) x US-09-702-572-1 (1-2920)

Qy 30 LeuLeuLeuAsnSerCysPheLeuLeuSerSerPheAsnGlyThr----- 44
Db 182 CTGCTGCTG-----TGCTGCTGTCAGTTCGTGCTTGGGCTCCGTCCTCCACG 235
Qy 45 -----AspLeuGluLeuArgLeuValAsn---GlyAsp 54
Db 236 GCCCTGAGAGAGCGCGGAGCGAGGCTTCGGTCCGCTGGCTGGCTTCCCCAGG 295
Qy 55 GlyProCysSerGlyThrValGluValLysPheGlnGlyGlnTrpGlyThrValCysAsp 74
Db 296 AAGCCCTACGAGGCGCGGTGAGATACAGCGAGCTGCTGAATGGGCGACCATCTCGCAT 355
Qy 75 AspGlyTrpAsnThrThrAlaSerThrValValCysLysGlnLeuLeuGlyCysProPheSer 94
Db 356 GATGACTTCACGCTCAGGCTGCCACATCTCTCCGCGGAGCTGGGC-----PTCACA 409
Qy 95 PheAla-----MetPheArgPheGlyGlnAlaValThrArgHisGlyLys 109
Db 410 GAGGCCACAGGCTGACCCACAGTCACCAATATGACCCCTGGNACA-----GGCCGC 460
Qy 110 IleTrpLeuAspAspValSerCysTyrGlyAsnGluSerAlaLeuTrpGluCysGlnHis 129
Db 461 ATCTGGCTGGACAACCTGAGTGCAGTGGGACCGAGCAGAGTGTGACTGAATGTGCTCC 520
Qy 130 ArgGluTrpGlySerHisAsnCysTyrHisGlyGluAspValGlyValAsnCys----- 147
Db 521 CGGGCTGGGGAGACAGTACGTACGACGATGAGGATGCTGGGGTCTCTCTGCAAGAAC 580
Qy 148 -----TyrGlyGluAlaAsnLeu----- 153
Db 581 CAGCGCTCCCTGGCTTCTCGACTCCAAATGTCATTGAGGTAGAGCATCACCTGCAAGTG 640
Qy 154 -----GlyLeuArgLeuValAspGlyAsnAsn 162
Db 641 GAGGAGGTGCGAATTCGACCCCGCTTGGTGGGCGACAGCACCCCTGCC----- 691
Qy 163 SerCysSerGlyArgValGluValLysPheGlnGluArgTrpGlyThrIleCysAspAsp 182
Db 692 GTGACGGAGGGCTGGTGAAGTACAGCTTCTGACGGCTGGTCCCAAGTGTGCGACAAA 751
Qy 183 GlyTrpAsnLeuAsnThrAlaAlaValValCysArgGlnLeuLeuGlyCysProSer---Ser 201
Db 752 GGTGGAGCGCCACACAGCCACGCTGCTGCGGGATGCTGGGCTTCCCGACGCAAAAG 811
Qy 202 PheIleSerSerGlyValValAsn-----SerProAlaValLeuArgProIleTrp-Le 219
Db 812 AGGGTCAACGC--GGCTTTCTACAGCTGCTAGCCCAACAGCGGACGACACTCTTTGGTCT 870
Qy 219 uAspAspIleLeuCysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTr 239
Db 871 GCATGGGTGGGTGGCGGACCGAGGCCACCTCTCCCTCTCTTCTCCCTGGAGTTCTA 930
Qy 239 pGlyAsnHisAsp-----CysSerHisAsnGluAspValThrLeuThrCys----- 254
Db 931 TCGTCCCAATGACACCCGCGAGTGCCTTGGGGGGGCGCTCGAGTGGTGGCTGTGTGCC 990
```

```

Qy 255 -----TyrAspSerSer----- 258
Db 991 AGGCCTCTGTACGGCGCATCCAGTGGCGAGAGAGCAACAGCTGAAGCCTCAGGG 1050
Qy 259 -AspLeuGluLeuArgLeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLy 278
Db 1051 GGAGGCCCTGTCTCCGCTCTAAAGGGCGGCGCCACCTGGAGAGGCGCGGTAGAAGTCT 1110
Qy 278 sIleGlnGlyArgTrpGlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspVa 298
Db 1111 GAAGCCACGACATGGGCACAGCTCTGTACCGCAAGTGGGACCTGCATGCACCCAGCT 1170
Qy 298 lValCysLysGlnLeuGlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuG 318
Db 1171 GGTGTGTCGGGAGCTGGCTTCGGGAGTCTCGAGAAGCTCTGAGTGGCGCTCGCATGG 1230
Qy 318 nSerGlySerAspValValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLe 338
Db 1231 GCAGGCGATGGTGTATCCACCTGAGTCAAGTTCGCTCTGCACAGGAGCTCTCCCT 1290
Qy 338 uTrpAspCysArgHisSerGlyThrValAsnPheAspCysLeuHisGlnAsnAspValSe 358
Db 1291 CTGGAAGTGGCCCCACAGAACATCACAGCTGAGATTGTTCCATAGCCAGGATGCCGG 1350
Qy 358 rValIleCysSer-----AspGlyAlaAspLeuGluLeuArgLeuAlaAspLySe 375
Db 1351 GGTCCGGTGCACACCTTACACTGGGCGAGAGACAGGATCCGACTCAGTGGGCGCG 1410
Qy 375 rAsnAsnCysSerGlyArgValGluValArgIleHisGlu-----GlnTrpTr 391
Db 1411 CAGCAACATGAGGGCGAGTGGAGTGCATAATAGGGGACCTGGGCGCTTCCCTGGGG 1470
Qy 391 pThrIleCysAspGlnAsnTrpLysAsnGlnAlaLeuValValCysLysGlnLeuG 411
Db 1471 CCTCATCTGGGGATGACTGGGCGACCTGGAGGCCATGGTGGCTGTAGGCAACTGG 1530
Qy 411 yCysProPheSerValPheGlySerArgAla-----LysProSerAsnGluAlaAr 429
Db 1531 TCTGGCTACGCCAACACACCGCTGCAGGAGACCTGGTACTGGGACTCTGGGAATATAAC 1590
Qy 429 gAspIleTrpIleAsnSerIleSerCysThrGlyAsnGluSerAlaLeuTrpAsnCysTh 449
Db 1591 AGAGTGTGTGAGTGGAGTGGCTGCACAGGAGTACTGAGTGTCCCTGGATCAGTGTGC 1650
Qy 449 rTyrAspGlyLysAlaLysArgThrCysPheArg-----ArgSerAspAlaGlyVa 466
Db 1651 CCATCATGCG---ACCCACATCACCTGCAAGAGGACAGGACCCGCTTCACTGCTGGAGT 1707
Qy 466 lIleCysSerAspLysAlaAspLeuArgLeuValGlyAlaHisSerProCystTy 486
Db 1708 CATCTGTCTGAGACTGCATCA---GATCTGTGTCTG---CACTCA----- 1747
Qy 486 rGlyArgLeuGluValLysTyrGlnGlyGluTrpGlyThrValCysHisAspArgTrpSe 506
Db 1748 -----CACTGTGTGCA--GGAGCCGCTTCA 1772
Qy 506 rThrArgAsnAla-----AlaValValCysLysGlnLeuGlyCysGlyLysProMe 523
Db 1773 TCGAAGACCGGCCCTGTCATATGTTGTACTGTGC-----TGGGGAAGAGAACT 1820
Qy 523 t-----HisValPheGlyMetThrThrPheLysGluAlaSerGlyProIle 538
Db 1821 GCCTGGCCAGCTACGCCGCTCAGCAACTGGCCCTATGGTCACCGGCTGTGCTCCGAT 1880
Qy 538 eTrpLeuAspAspVal-SerCys-----lIleGlyAsnGlu 550
Db 1881 TCTCTCCCATGATCCACACCTGGGACGAGCTGACTTCAAGGCCCAAGGCTGGGCGCCACT 1940
Qy 550 eRAsnIleTrp---AspCysGluHisSerGlyTrpGlyLysHisAsnCysValHisArgG 569
Db 1941 CTGGGTGTGGCACGAGTGC-----CATGGCATATTACCACAGCA 1979
```

```
Qy 569 luAspVal-----lleValThrCysSerGlyAspAlaThrTrpGlyLeuA 584
Db 1380 TGGACATCTTCACACTATGATATCTCACCCCAATGGCAACCAAGGTG----- 2029
Qy 584 rgLeuValGlyGlySerAsnArgCysSerGlyArgLeuGluValTyPheGlnGlyArgT 604
Db 2030 -----GCTGAGGGCCACA 2042
Qy 604 rpGlyThrValCysAspAspGlyTrpAsnSerLysAlaAlaValValCysSerGlnL 624
Db 2043 AAGCTAGTTCTGT----- 2056
Qy 624 euAspCysProSerSerIleGlyMetGlyLeuGlyAsnAlaSerThrGlyTyrglyL 644
Db 2056 ----- 2056
Qy 644 ysIleTrpLeuAspAspValSerCysAspGlyAspGluSerAspLeuTrpSerCysArg 664
Db 2057 -----CTCGAAGACACTGAGTCTCAGGAGGATGCTCCAACGGGTATGAGTGCCCA 2108
Qy 664 snSerGlyTrpGlyAsnAsnAspCysSerHisSerGluAspValGlyValIleCysSerA 684
Db 2109 ACTTTGA-----GAGCAAGGCATCAGTGTGGTGTCTGGG 2144
Qy 684 sp-AlaSerAspMetGluLeuArgLeuValGlyGlySerSerArg 698
Db 2145 ATCTCTACCGCATGACATTCAGTGTCTCAGTGTGATGATGATCAGCG 2189
RESULT 12
US-09-341-587-7
; Sequence 7, Application US/09341587
; Patent No. 6346606
; GENERAL INFORMATION:
; APPLICANT: Mollenhauer, Jan
; TITLE OF INVENTION: Protein Containing an SRCR Domain
; FILE REFERENCE: 4121-108
; CURRENT APPLICATION NUMBER: US/09/341,587
; EARLIER FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: PCT/DE98/00096
; EARLIER FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 28720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-341-587-7
Alignment Scores:
Pred. No.: 1.18e-45 Length: 28720
Score: 618.50 Matches: 386
Percent Similarity: 33.63% Conservative: 183
Best Local Similarity: 22.81% Mismatches: 501
Query Match: 7.60% Indels: 629
DB: 4 Gaps: 80
US-09-759-130B-381 (1-1453) x US-09-341-587-7 (1-28720)
Qy 25 valValThrCysIleLeuLeuLeuAsnSerCysPheLeuIleSerSerPheAsnGly---- 43
Db 2872 ATTTGTCACCTCCCTGGTGGGATGATGAAGGTCTTTGTTCTCCCTGTAGGATCTGAA 2931
Qy 44 ThrAspLeuGluLeuArgLeuValAsnGlyAspGlyProCysSerGlyThrValGluVal 63
Db 2932 TCCACTTTGGCCCTGAGACTGGTGAATGGAGGTGACAGGTGTCAGAGCCGAGTGGAGTGC 2991
Qy 64 LysPheGlnGlyIntrpGlyThrValCysAspAspGlyTrpAsnThrThrAlaSerThr 83
Db 2992 CTATACCAAGGCTCTCTGGGGCCCGCTGTGTGATGACTACTGGACACCAATGATGCCAAC 3051
Qy 84 ValValCysLysGlnLeuGlyCysProPheSerPheAlaMet-----PheArgPhe 100
Db 3052 GTGGTCTGACGACGCTGGGCTGTGGTGGGCCATGTGTCAGGCCCAAGAAATGCCAGTTT 3111
```

```
Qy 101 GlyGlnAlaValThrArgHisGlyLysIleTrpLeuAspAspValSerCysTyrglyAsn 120
Db 3112 GGCAGAGGC-----TCAGGAGCCCATGCTCTGGATGATGCGCTGCTCAGGACAC 3162
Qy 121 GluSerAlaLeuTrpGluCysGlnHisArgGluTrpGlySerHisAsnCysTyrglyHisGly 140
Db 3163 GAGTCTTACCTGTGGAGTGGCCGCCACAAATGGCTCTGCCACAAATGTTGGCATCAT 3222
Qy 141 GluAspValGlyValAsnCysTyrgly-----GluAlaAsnLeuGly 154
Db 3223 GAAGATGCTGTGTGTCATCTGCTCAGGTGGGCTTTCAAGACCTTGGGCTCCCTCTCTTAAG 3282
Qy 155 LeuArgLeuValAspGlyAsnAsn----- 162
Db 3283 TTGAAGTTGCTCAGGAAGAAATCCTAATTACATTCCTGATCTCTCACTCAAAAGCTTTT 3342
Qy 163 -----SerCysSerGlyArgValGluValLysPheGlnGlu 174
Db 3343 TCTATGTTTCTATATTTCTGAAGTCTTGTAGCTCTCTGCTAAGAAATCTTTATGAATTT 3402
Qy 175 ArgTrpGlyThrIleCysAspAspGlyTrpAsnLeuAsnThrAlaAlaValValCysArg 194
Db 3403 TGCTACAGTACCTGGTGCAGC-----TGTGGC 3429
Qy 195 GlnLeuGlyCysProSerSerPheIleSerSerGlyValValAsnSerProAla----- 212
Db 3430 CACTTAGCCAGGGCTCCGAACCTGAA-ACAACAACCCAGACTTATCCCTCCCTCCTGAGGC 3488
Qy 213 -----valLeuArgProIleTrpLeuAspIleLeu 223
Db 3489 AGTGAAGGAAGAGGAGAGAGAGAAAGTGTGGCTCCC----- 3527
Qy 224 CysGlnGlyAsnGluLeuAlaLeuTrpAsnCysArgHisArgGlyTrpGlyAsnHisAsp 243
Db 3528 ---CACGGCTCCATTTCTTCCCTGCTGAGTAGCAGCTGTTGAGGGTATCTGTGGACACAGC 3584
Qy 244 -----CysSerHisAsnGluAspVal---ThrLeuThrCys 254
Db 3585 ACAGATAGCAGGGGAGGGAGGGATCTCTCCTACACAGGAACCTGTGAACATAA-AGATGC 3643
Qy 255 TyrAspSerSer-----AspLeuGluLeuArg 263
Db 3644 TTGCTGGAAGTGGTCTTCTCAGCTGAGACCCAGGAGGTCTGGAATAAGAGCTCAA 3703
Qy 264 LeuValGlyGlyThrAsnArgCysMetGlyArgValGluLeuLysIleGlnGlyArgTrp 283
Db 3704 GGGTTAGAGT-GCAATGTCATGT-----CTGTTGTCTCAGGCTGCTT 3747
Qy 284 GlyThrValCysHisHisLysTrpAsnAsnAlaAlaAspValValCysLysGlnLeu 303
Db 3748 GGAGTCTCTGACCTCAGGTCTCTTAAGAATGCTGCAGAGCAGCTGCTGTGCC-CAGGTC 3806
Qy 304 GlyCysGlyThrAlaLeuHisPheAlaGlyLeuProHisLeuGlnSerGlySerAspVal 323
Db 3807 TGGTGTGGG-----GAGGCAGCCCCCAT---GAGTCTGGCCAG----- 3842
Qy 324 ValTrpLeuAspGlyValSerCysSerGlyAsnGluSerPheLeuTrpAspCys----- 341
Db 3843 -----GCATGGCTTGTATTATGTC 3860
Qy 342 -----ArgHis----- 343
Db 3861 CTGTGGTGGGGCTGTAGATCACACAGGCAATTTGGGCTGGAGTGGCTCTCTCAGCCTT 3920
Qy 344 -----SerGlyThr-----ValAsnPheAsp 350
Db 3921 GCTGACTCAGGAACCTGCCAAAACATCCAAAGGAGAGTCTTGGTTTGGGTCAACCTGGTC 3980
Qy 351 CysLeuHisGlnAsnAspValSerValIleCysSerAsp-----GlyAlaAspLeuGlu 368
Db 3981 ACCCTGGCAGACACAAAGTTACTCACCTCGAGCTGACAATAGTGGGCGAGATCTGCT 4040
```



Db 5738 GTCACAGG-----TGCTTCCCAATACTTGTAGCTTCCATAGACTTGGGTGGAGTAGGAC 5791  
Qy 978 --AsnCys-GlnMetThrValLeuGlyAlaProProCysIleHisGlyAsnThrValSer 996  
Db 5792 ATCACTGCTTCTCCATATCATCAAGCTGACCTCTGTCTGTATTCATATTCAAAGGTG 5851  
Qy 997 ValIleCysThr-----GlySerLeuThrGln-Pro----- 1006  
Db 5852 ATTACCTGCACACAGCTGCTTGTGGCAATTAGGAAGTGCCTGTAGTGTGAACATTCCTTTA 5911  
Qy 1007 -LeuPheProCysLeuAlaAsnValSerAspProTyrIleuSerAlaValProGluGlySe 1026  
Db 5912 AATCCTTGACCTCATATCATGATGGAAGGCTTCTGTGTCTCC---CCTGTAGGATC 5968  
Qy 1026 rAlaLeuIleCysLeuGluAspLysArgLeuArgLeuValAspGlyAspSerArgCysAl 1046  
Db 5969 TGAA-----TCCAGTTTGGCTCTGAGGCTGGTGAATGGAGGTGACAGGTGTCG 6016  
Qy 1046 aGlyArgValGluIleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAs 1066  
Db 6017 AGGCCGAGTGGAGGTCTGTATCGAGGCTCTGGGAACCGTGTGTATGACAGCTGGGA 6076  
Qy 1066 pLeuSerAspAlaHisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaTh 1086  
Db 6077 CACCAATGATGCCAATGGTCTGCAGCAGCTGGGCTGTGGCCATGTGCGGCC 6136  
Qy 1086 rValSerAlaHisPheGlyGluGlySerGlyProIleTrpLeuAspLeuAsnCysTh 1106  
Db 6137 AGGAATGCCGGTGTGGCAGGCTCAGGACCATTCCTCTGGATGATGTGCGCTGCTC 6196  
Qy 1106 rGlyThrGluSerHisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysAr 1126  
Db 6197 AGGGAATGAGTCTACCTGTGGAGCTGCCCCACAAAGGCTGGCTCACCCACAACTGGG 6256  
Qy 1126 gHisGlyAspAlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGl 1146  
Db 6257 CCATCAGCAAGACGCTGGTGTCTCTGCTCA----- 6287  
Qy 1146 uThrGluThrGluSerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySe 1166  
Db 6288 -----GGTGGGCC--TTCAGACCTGGGGCTCCCTCTCTGGGGTGG 6327  
Qy 1166 rVal---GlyArgArgAsn-----lleThrAlaIleAlaGl 1178  
Db 6328 AGTTTGTCTCCAGAAACCTCTAATTACATTCCTCTCTCTCACTCAAGCTTCTCTCGT 6387  
Qy 1178 yIleValCys----- 1181  
Db 6388 TGTTTCTCTGCTTTTGAAGACTTGTAGCTCTCTGTGAAGATCCATATGAATTCACGT 6447  
Qy 1182 -----ArgGlnLeuGlyCysGlyGluAsnGlyValValSer----- 1193  
Db 6448 CCTAGTGTCTCTGGTCTAGGACAGGGGACCAACTCAACACACACCCAGACTTT 6507  
Qy 1194 -----LeuAlaPro----- 1196  
Db 6508 ATCCCTTCTGTAGCAGTCAAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6567  
Qy 1197 -----LeuSerLysThrGlySerGlyPheMetTrpValAspAspIleGl 1211  
Db 6568 CTCCATTTCTCCCTACTGAGTACGCTGGGTGAGGTATGG--TGGACACAGCAGACAGA 6626  
Qy 1211 nCysProLysThrHisIle-----SerIleTrpGlnCysLeuSerAlaPr 1226  
Db 6627 CGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6686  
Qy 1226 oTrpGluArgArgIleSer-----SerProAlaGluGluThrTrpIleThrCysGluAs 1244  
Db 6687 TTGGGTAGGGAATTCCTACTTCTAGGCCCCAGTAGGAGGAGGAGGAGGAGGAGGAGGAG 6735  
Qy 1244 pArgIleArgValArgGlyGlyAspThrGluCysSerGlyArgValGluIleTrpHisAl 1264  
Db 6736 -----GGAGGC-----ATATGGGCTAA 6752

Qy 1264 aGlySerTrpGlyThrValCysAspSerTrpAspLeuAlaGluAlaGluValValCy 1284  
Db 6753 GAGTGCACACGGGTATCTGTGC-----ATGTG 6779  
Qy 1284 sGlnGlnLeuGlyCysGlySerAlaLeuAlaLeuArgAspAlaSerPheGlyGlnGl 1304  
Db 6780 TGAC-----CTGGG 6788  
Qy 1304 yThrGlyThrIleTrpLeuAspMetArgCysLysGlyAsn-----GluSerPh 1321  
Db 6789 TCTTGGTTCAGTTTCAGCTCTCTCCAAAGAACTGCAGTGCATGTGTCTAGCAGGATG 6848  
Qy 1321 eLeuTrpAsp-----CysHisAlaLysPro-----Tr 1330  
Db 6849 GTGTTGGGAGGCGAGCTCCCATGAGATCTGCAGGCAAGGCTTGTATTACGTGTGTT 6908  
Qy 1330 pGlyGlnSerAspCysGlyHisLysGluAsp-----AlaGly----- 1342  
Db 6909 TGGGATGGAAGT-----CACACAGGGGATTTTGTCTGGAGTGTGTCTGAGACTTGC 6962  
Qy 1343 -----ValArgCys-----SerGlyGlnSerLeuLysSerLeuAsnAl 1355  
Db 6963 TGACTTGTGTGAATTGCTAAACCTGCAAGGAAAGGTTGTTTGTTCATTCAATTGGACAT 7022  
Qy 1355 aSerSerGlyHis----- 1361  
Db 7023 CACATCAGACACAGAGTTAATTCCTTGGAGCTGGAATGTTGGACAGAAATCTGCTTGG 7082  
Qy 1361 aLeuIleLeuSerSerIlePhe 1368  
Db 7083 ATCCCTTATAAGGAGCGTCTTT 7104  
RESULT 13  
US-08-477-674-9  
; Sequence 9, Application US/08477674  
; Patent No. 5644035  
; GENERAL INFORMATION:  
; APPLICANT: Koths, Kirston E.  
; APPLICANT: Halenbeck, Robert F.  
; APPLICANT: Taylor, Eric W.  
; APPLICANT: Wang, Alice M.  
; APPLICANT: Caspitt, Clayton L.  
; TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cetus Oncology Corporation  
; STREET: 1400 Fifty-Third Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,674  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/316,714  
; FILING DATE:  
; APPLICATION NUMBER: US/07/961,404  
; FILING DATE: 15-OCT-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Kenneth M.  
; REGISTRATION NUMBER: 34,174  
; REFERENCE/DOCKET NUMBER: 2595.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 420-3152  
; TELEFAX: (510) 658-5470

TELEX: N/A  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2285 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-477-674-9

Alignment Scores:  
Pred. No.: 3,56e-19 Length: 2285  
Score: 314.00 Matches: 107  
Percent Similarity: 42.00% Conservatives: 107  
Best Local Similarity: 30.57% Mismatches: 108  
Query Match: 3.86% Indels: 96  
DB: 1 Gaps: 15

US-09-759-130B-381 (1-1453) x US-08-477-674-9 (1-2285)

QY 1031 LeuGluAspLysArgLeuValAspGlyAspSerArgCysAlaGlyValGlu 1050  
DB 234 GTGAATGATGGTACATGGCTGGCGATGGGGCGCCACCAACAGGCGCGTGGAG 293  
QY 1051 IleTyrHisAspGlyPheTrpGlyThrIleCysAspAspGlyTrpAspLeuSerAspAla 1070  
DB 294 ACTTCTCAGAGGCGAGTGGGGGCTGTGTGTGACACCTCTGGGACCTGACTGATGCC 353  
QY 1071 HisValValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrValSerAlaHis 1090  
DB 354 AGCGTGTCTGGCGGCTGGGCTTCGAGACGCCACCCAGGCTCTGGCGAGCTGCC 413  
QY 1091 PheGlyGluGlySerGlyProIleTrpLeuAspLeuAsnCysThrGlyThrGluSer 1110  
DB 414 TTGGGCAAGGATCAGGCCCATCATGCTGTGGACGAGTCCAGTCCAGCGGAAACCGAGGCC 473  
QY 1111 HisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAsp 1130  
DB 474 TCATGGCGGCTGCAAGTCCCTGGCTGGTGAAGACCACTGAGCGACGAGAGAGAC 533  
QY 1131 AlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGlu 1150  
DB 534 GCTGGTGTGTCTGCACCAATGAAACC-----AGGAGCACCCACACCCCTG 578  
QY 1151 SerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArg 1170  
DB 579 GACCTCTCC-----AGG 590  
QY 1171 AsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGly 1190  
DB 591 GAGCTCTCGAGGCGCTTGGCGAGATCTTTGACAGCCAGCGGGCTGGCACTGTCC----- 647  
QY 1191 ValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAlle 1210  
DB 648 ---ATCAGCGTGAATGTGCAGGCGAGGAGCGCTGGCTTC----- 686  
QY 1211 GlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArg 1230  
DB 687 ---TGTGGCCACACGTCATC-----CTGACTGCCAACCTGGAG----- 722  
QY 1231 IleSerSerProAlaGluThrTrpIleThrCysGluAspArgIleArgValArgGly 1250  
DB 723 -----GCCAGCGCCCTGTGGAGAGGAGCGCGGAGCAATGTACCATG---AGT 767  
QY 1251 GlyAspThrGluCys-----SerGlyArg 1258  
DB 768 GTGGATGCTGAGTGTGTGCCCATGTGTGAGGAGGACCTTCTCAGTACTTCTACTCCCGAAG 827  
QY 1259 ValGluIle-----TrpHisAlaGlySerTrpTrpGlyThrValCys 1271  
DB 828 ATTGACATCACCCCTGTGCTGCTCAAGTCTTCCAC----- 863  
QY 1272 AspAspSerTrpAspLeuAlaGluValValCysGlnGlnLeuGly---CysGly 1290

DB 864 -----AGCTGGCCTCTGCCTATGGGGCCAGCGAGCTGCTGAGGCTACTCGCA 911  
QY 1291 SerAlaLeuAlaLeu-----ArgAspAlaSerPheGlyGlnGlyThrGlyThrIle 1308  
DB 912 AGCCTCTTTGCCATCTCTCTCCCGCAGGACCCCTGTTCCA-GATGCCCTGGACCTGTA 970  
QY 1309 TrpLeuAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLys 1328  
DB 971 TGCCTA-----TGCAGTGGCCAC-----AGGGGACGC 997  
QY 1329 ProTrpGlyGlnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGln 1348  
DB 998 CCTGTGGA-----GAAGCTCTGCCTACAGTTCTCTGCGCTG 1033  
QY 1349 SerLeuLysSerLeuAsnAlaSerSerGly 1358  
DB 1034 GAACCTCGAGGCTTCAGCAGCGCCGAGGC 1063

RESULT 14  
US-08-473-791-9  
Sequence 9, Application US/08473791.  
Patent No. 5736340  
GENERAL INFORMATION:  
APPLICANT: Koths, Kirston E.  
APPLICANT: Halenbeck, Robert F.  
APPLICANT: Taylor, Eric W.  
APPLICANT: Wang, Alice M.  
APPLICANT: Caspit, Clayton L.  
TITLE OF INVENTION: Secreted Mac-2-Binding Glycoprotein  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cetus Oncology Corporation  
STREET: 1400 Fifty-Third Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,791  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/316,714  
FILING DATE:  
APPLICATION NUMBER: US/07/961,404  
FILING DATE: 15-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Kenneth M.  
REGISTRATION NUMBER: 34,174  
REFERENCE/DOCKET NUMBER: 2595.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 420-3152  
TELEFAX: (510) 658-5470  
TELEX: N/A  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2285 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-473-791-9

Alignment Scores:  
Pred. No.: 3,56e-19 Length: 2285  
Score: 314.00 Matches: 107  
Percent Similarity: 42.00% Conservatives: 40

DB 998 CCTGCTGGA-----GAAGCTCTGCGCTACAGTTCTCTGGCGCTG 1039

DB:	2	Gaps:	15
US-09-759-130B-381 (1-1453) x US-08-316-714-9 (1-2285)			

QY 1071 HisValCysGlnLysLeuGlyCysGlyValAlaPheAsnAlaThrVal

Db 354 AGCGTCTGTCGGCGCCCTGGGCTTCGAGAACGCCACCGAGGCTCTGGGAGAGCTGCC 413  
QY 1091 PheGlyGluGlySerGlyProIleTrpLeuAspAspLeuAsnCysThrGlyThrGluSer 1110  
Db 414 TTCGGGCAAGGATCAGGCCCATCATGCTGGACGAGGTCCAGTGCACGGGACCGAGGCC 473  
QY 1111 HisLeuTrpGlnCysProSerArgGlyTrpGlyGlnHisAspCysArgHisLysGluAsp 1130  
Db 474 TCACCTGGCGGACTGCAAGTCCCTGGCTGGCTGAAGAGCACTGACGAGGACGAGAGAC 533  
QY 1131 AlaGlyValIleCysSerGluPheThrAlaLeuArgLeuTyrSerGluThrGluThrGlu 1150  
Db 534 GCTGTGTGGTCTGCACCAATGAAACC-----AGGAGCACCCACACACCTGTG 578  
QY 1151 SerCysAlaGlyArgLeuGluValPheTyrAsnGlyThrTrpGlySerValGlyArgArg 1170  
Db 579 GACCTCTCC-----AGG 590  
QY 1171 AsnIleThrThrAlaIleAlaGlyIleValCysArgGlnLeuGlyCysGlyGluAsnGly 1190  
Db 591 GAGCTCTCGAGGCGCTTGGCCAGATCTTTGACAGCCAGCGGGGCTGGACCTGTCC--- 647  
QY 1191 ValValSerLeuAlaProLeuSerLysThrGlySerGlyPheMetTrpValAspAspIle 1210  
Db 648 ---ATCAGCGTGAATGTCGAGGCGGAGGAGCGCCCTGGGCTTC----- 686  
QY 1211 GlnCysProLysThrHisIleSerIleTrpGlnCysLeuSerAlaProTrpGluArgArg 1230  
Db 687 ---TGTGGCCACACGTCATC-----CTGACTGCCAACCTGGAG----- 722  
QY 1231 IleSerSerProAlaGluThrTrpIleThrCysGluAspArgIleArgValArgGly 1250  
Db 723 -----GCCAGGCGCTGTGGAAGAGCGCGGCGCAATGTCACCATG---AGT 767  
QY 1251 GlyAspThrGluCys-----SerGlyArg 1258  
Db 768 GTGGATGCTGAGTGTGTGCCCATGTCAGGGACCTTCTCAGGTACTTCTACTCCCGAAGG 827  
QY 1259 ValGluIle-----TrpHisAlaGlySerTrpGlyThrValCys 1271  
Db 828 ATTGACATCACCCCTGTCGTCAGTCAAGTGTCTCCAC----- 863  
QY 1272 AspAspSerTrpAspLeuAlaGluAlaGluValValCysGlnGlnLeuGly---CysGly 1290  
Db 864 -----AAGCTGGCCTCTGCCTATGGGGCCAGGAGCTGCAGGGCTACTGCCGA 911  
QY 1291 SerAlaLeuAlaLeu-----ArgAspAlaSerPheGlyGlnGlyThrGlyThrIle 1308  
Db 912 AGCCTCTTTGCCATCTCTCTCCCGGAGACCCCTGTTCCA-GATGCCCTCGACCTGTA 970  
QY 1309 TrpLeuAspAspMetArgCysLysGlyAsnGluSerPheLeuTrpAspCysHisAlaLys 1328  
Db 971 TGCCTA-----TGCAGTGGCCAC-----AGGGAGCGC 997  
QY 1329 ProTrpGlyClnSerAspCysGlyHisLysGluAspAlaGlyValArgCysSerGlyGln 1348  
Db 998 CCTGCTGGA-----GAAGCTCTGCTACACTTCTGSCCTG 1033  
QY 1349 SerLeuLysSerLeuAsnAlaSerSerGly 1358  
Db 1034 GAACCTCGAGGCTTGACGAGCGGCGGAGGC 1063

Search completed: May 12, 2003, 19:17:06  
Job time : 236 secs